2

Exact Simulation of Solutions of SDEs

Accurate scenario simulation methods for solutions of multi-dimensional stochastic differential equations find applications in the statistics of stochastic processes and many applied areas, in particular in finance. They play a crucial role when used in standard models in various areas. These models often dominate the communication and thinking in a particular field of application, even though they may be too simple for advanced tasks. Various simulation techniques have been developed over the years. However, the simulation of solutions of some stochastic differential equations can still be problematic. Therefore, it is valuable to identify multi-dimensional stochastic differential equations with solutions that can be simulated exactly. This avoids several of the theoretical and practical problems of those simulation methods that use discrete-time approximations. This chapter follows closely Platen & Rendek (2009a) and provides methods for the exact simulation of paths of multi-dimensional solutions of stochastic differential equations, including Ornstein-Uhlenbeck, square root, squared Bessel, Wishart and Lévy type processes. Other papers that could be considered to be related with exact simulation include Lewis & Shedler (1979), Beskos & Roberts (2005), Broadie & Kaya (2006), Kahl & Jäckel (2006), Smith (2007), Andersen (2008), Burq & Jones (2008) and Chen (2008).

2.1 Motivation of Exact Simulation

Avoiding any error in the simulation of the path of a given process can only be achieved in exceptional cases. However, when it is possible it makes the numerical results very accurate and reliable.

Accurate scenario simulation of solutions of SDEs is widely applicable in stochastic analysis itself and many applied areas, in particular, in quantitative finance and for dynamic financial analysis in insurance, see Kaufmann, Gadmer & Klett (2001). Monographs in this direction include, for instance, Kloeden & Platen (1999), Kloeden, Platen & Schurz (2003), Milstein

E. Platen, N. Bruti-Liberati, *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, Stochastic Modelling and Applied Probability 64, DOI 10.1007/978-3-642-13694-8_2,

(1995a), Jäckel (2002) and Glasserman (2004). Many discrete-time simulation methods have been developed over the years. However, some SDEs can be problematic in terms of discrete-time approximation via simulation. Therefore, it is necessary to understand and avoid the problems that may arise during the simulation of solutions of such SDEs. For illustration, let us consider a family of SDEs of the form

$$dX_t = a(X_t)dt + \sqrt{2X_t}dW_t \tag{2.1.1}$$

with some given drift coefficient function $a(x)$. Note that the diffusion coefficient function $f(x) = \sqrt{2x}$ is here non-Lipschitz. Its derivative becomes infinite as x tends to 0. The standard convergence theorems, derived in the previously mentioned literature and presented later in this book, do not easily cover such cases. It is, therefore, of interest to identify approximate simulation methods for various types of nonlinear SDEs and also for multi-dimensional SDEs. We will emphasize the fact that the problem of non-Lipschitz coefficients is circumvented for SDEs where we can simulate exact solutions. For squared Bessel processes of integer dimension, see Revuz & Yor (1999) and Platen & Heath (2006), we will explain how to simulate such solutions. Exact solutions can be simulated for a range of diffusion processes by sampling from their explicitly available transition density for some special cases of nonlinear SDEs where the drift function $a(\cdot)$ in (2.1.1) takes a particular form, see Craddock & Platen (2004). These will also include squared Bessel processes of noninteger dimensions, see Sect. 2.2.

Another problem with the simulation of SDEs may be the lack of sufficient numerical stability of the chosen scheme. As we will discuss later in Chap. 14, numerical stability is understood as the ability of a scheme to control the propagation of initial and roundoff errors. Numerical stability may be lost for some parameter ranges of a given SDE when using certain simulation schemes with a particular time step size. The issue of numerical stability can be circumvented when it is possible to simulate exact solutions.

Moreover, for theoretically strictly positive processes it is often not sufficient to use simulation methods that may generate negative values. This problem however can, in some cases, be solved by a transformation of the initial SDE, by use of the Itô formula, to a process which lives on the entire real axis. This is, in particular, useful for geometric Brownian motion, the dynamics of the Black-Scholes model, where one can take the logarithm to obtain a linearly transformed Wiener process. One may try such an approach to transform the *square root process* of the form

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \tag{2.1.2}$$

where $t \in \Re^+$. This process remains strictly positive for dimension $\delta = \frac{4\kappa\theta}{\sigma^2} > 2$. Suppose that we simulate for $\delta > 2$ the process $Y_t = \sqrt{X_t}$ using a standard explicit numerical scheme such as the Euler scheme, see Kloeden & Platen (1999). The SDE of the corresponding stochastic process $Y = \{Y_t = \sqrt{X_t}, t \geq 0\}$ has then additive noise and is given by

$$dY_t = \left(\frac{\kappa\theta - \sigma^2/4}{2Y_t} - \frac{\kappa}{2} Y_t \right) dt + \frac{\sigma}{2} dW_t. \tag{2.1.3}$$

Theoretically, by squaring the resulting trajectory of Y we should obtain an approximate trajectory of the square root process X. However, note that the drift coefficient $\frac{\kappa\theta - \sigma^2/4}{2y} - \frac{\kappa}{2} y$ is non-Lipschitz and may almost explode for small y. Even though we have additive noise this feature will most likely produce simulation problems near zero. This kind of problem becomes even more serious for small dimension $\delta < 2$ of the square root process. It would be very valuable to have an exact solution, which avoids this kind of problems.

After the Wiener process and its direct transformations, including geometric Brownian motion and the Ornstein-Uhlenbeck process, the family of square root and squared Bessel processes are probably the most frequently used diffusion models in applications. In general, it is a challenging task to obtain, efficiently, a reasonably accurate trajectory of a square root process using simulation, as is documented in an increasing literature on this topic. Here we refer to the work of Deelstra & Delbaen (1998), Diop (2003), Bossy & Diop (2004), Berkaoui, Bossy & Diop (2005), Alfonsi (2005), Broadie & Kaya (2006), Lord, Koekkoek & van Dijk (2006), Smith (2007) and Andersen (2008). We will also study the simulation of multi-dimensional square root and squared Bessel processes below.

In various areas of application of stochastic analysis one has to model vectors or even matrices of evolving dependent stochastic quantities. This is typically the case, for instance, when modeling related asset prices in a financial market. All the above mentioned numerical problems can arise in a complex manner when simulating the trajectories of such multi-dimensional models. For instance, different time scales in the dynamics of certain components can create stiff SDEs in the sense of Kloeden & Platen (1999), which are almost impossible to handle by standard discrete-time schemes. This makes it worthwhile to identify classes of multi-dimensional SDEs with exact solutions. In addition, we will see that almost exact approximations will be of particular interest.

2.2 Sampling from Transition Distributions

In the following we will consider some multi-dimensional diffusion processes with explicitly known transition distributions. Since, it is rare that one has an exact formula for a transition distribution, these examples are of particular interest. We will show how to use explicitly available multivariate transition distributions for the simulation of exact solutions of SDEs. In fact, for certain multi-dimensional diffusions, given by some system of SDEs, one may need extra information about their behavior at finite boundaries to obtain a complete description of the modeled dynamics. The description of the diffusions via transition distributions contains this information. One needs to keep this

in mind since most of the literature is primarily concerned with the modeling via SDEs.

Inverse Transform Method

The following well-known *inverse transform method* can be applied for the generation of a continuous random variable Y with given probability distribution function F_Y. From a uniformly distributed random variable $0 < U < 1$, we obtain an F_Y distributed random variable $y(U)$ by realizing that

$$U = F_Y(y(U)), \tag{2.2.1}$$

so that

$$y(U) = F_Y^{-1}(U). \tag{2.2.2}$$

Here F_Y^{-1} denotes the inverse function of F_Y. More generally, one can still set

$$y(U) = \inf\{y : U \leq F_Y(y)\} \tag{2.2.3}$$

in the case when F_Y is no longer continuous, where $\inf\{y : U \leq F_Y(y)\}$ denotes the lower limit of the set $\{y : U \leq F_Y(y)\}$. If U is a $U(0,1)$ random variable, then the random variable $y(U)$ in (2.2.2) will be F_Y-distributed. The above calculation in (2.2.2) may need to apply a root finding method, for instance, a Newton method, see Press, Teukolsky, Vetterling & Flannery (2002). Obviously, given an explicit transition distribution function for the solution of a one-dimensional SDE we can sample a trajectory directly from this transition law at given time instants. One simply starts with the initial value, generates the first increment and sequentially the subsequent random increments of the simulated trajectory, using the inverse transform method for the respective transition distributions that emerge.

Also in the case of a two-dimensional SDE we can simulate by sampling from the bivariate transition distribution. We first identify the marginal transition distribution function F_{Y_1} of the first component. Then we use the inverse transform method, as above, for the exact simulation of an outcome of the first component of the two-dimensional random variable based on its marginal distribution function. Afterwards, we exploit the conditional transition distribution function $F_{Y_2|Y_1}$ of the second component Y_2, given the simulated first component Y_1, and use again the inverse transform method to simulate also the second component of the considered SDE. This simulation method is exact as long as the root finding procedure involved can be interpreted as being exact. It exploits a well-known basic result on multi-variate distribution functions, see for instance Rao (1973).

It is obvious that this simulation technique can be generalized to the exact simulation of increments of solutions of some d-dimensional SDEs. Based on a given d-variate transition distribution function one needs to find the marginal distribution F_{Y_1} and the conditional distributions $F_{Y_2|Y_1}$, $F_{Y_3|Y_1,Y_2}$,

..., $F_{Y_d|Y_1,Y_2,\ldots,Y_{d-1}}$. Then the inverse transform method can be applied to each conditional transition distribution function one after the other. This also shows that it is sufficient to characterize explicitly in a model just the marginal and conditional transition distribution functions.

Note also that nonparametrically described transition distribution functions are sufficient for application of the inverse transform method. Of course, explicitly known parametric distributions are preferable for a number of practical reasons. They certainly reduce the complexity of the problem itself by splitting it into a sequence of problems. Further, we will list below important examples of explicitly known multi-variate transition densities and distribution functions.

Copulas

Each multi-variate distribution function has its, so called *copula*, which characterizes the dependence structure between the components. Roughly speaking, the copula is the joint density of the components when they are each transformed into $U(0,1)$ distributed random variables. Essentially, every multivariate distribution has a corresponding copula. Conversely, each copula can be used together with some given marginal distributions to obtain a corresponding multi-variate distribution function. This is a consequence of Sklar's theorem, see for instance McNeil et al. (2005).

If, for instance, $\boldsymbol{Y} \sim N_d(\boldsymbol{\mu}, \boldsymbol{\Omega})$ is a Gaussian random vector, then the copula of $\boldsymbol{Y}$ is the same as the copula of $\boldsymbol{X} \sim N_d(\mathbf{0}, \boldsymbol{\Omega})$, where $\mathbf{0}$ is the zero vector and $\boldsymbol{\Omega}$ is the correlation matrix of $\boldsymbol{Y}$. By the definition of the d-dimensional Gaussian copula we obtain

$$C_{\boldsymbol{\Omega}}^{Ga} = P(N(X_1) \leq u_1, \ldots, N(X_d) \leq u_d) = N_\Omega(N^{-1}(u_1), \ldots, N^{-1}(u_d)), \tag{2.2.4}$$

where N denotes the standard univariate normal distribution function and N_Ω denotes the joint distribution function of $\boldsymbol{X}$. Hence, in two dimensions we obtain

$$C_{\boldsymbol{\Omega}}^{Ga}(u_1, u_2) = \int_{-\infty}^{N^{-1}(u_1)} \int_{-\infty}^{N^{-1}(u_2)} \frac{1}{2\pi(1-\varrho^2)^{1/2}} \exp\left\{\frac{-(s_1^2 - 2\varrho s_1 s_2 + s_2^2)}{2(1-\varrho^2)}\right\} \times ds_1\, ds_2, \tag{2.2.5}$$

where $\varrho \in [-1, 1]$ is the correlation parameter in $\boldsymbol{\Omega}$.

Another example of a copula is the Clayton copula. This copula can be expressed in the d-dimensional case as

$$C_\theta^{Cl} = (u_1^{-\theta} + \ldots + u_d^{-\theta} - d + 1)^{-1/\theta}, \quad \theta \geq 0, \tag{2.2.6}$$

where the limiting case $\theta = 0$ is the d-dimensional independence copula. Moreover, d-dimensional Archimedian copulas can be expressed in terms of

Laplace-Stieltjes transforms of distribution functions on $\Re^+$. If F is a distribution function on $\Re^+$ satisfying $F(0) = 0$, then the Laplace-Stieltjes transform can be expressed by

$$\hat{F}(t) = \int_0^\infty e^{-tx} dF(x), \quad t \geq 0. \tag{2.2.7}$$

Using the Laplace-Stieltjes transform the d-dimensional Archimedian copula has the form

$$C^{Ar}(u_1, \ldots, u_d) = E\left(\exp\left\{-V \sum_{i=1}^{d} \hat{F}^{-1}(u_i)\right\}\right) \tag{2.2.8}$$

for strictly positive random variables V with Laplace-Stieltjes transform $\hat{F}$. A simulation method follows directly from this representation, see Marshall & Olkin (1988). More examples of multi-dimensional copulas can be found in McNeil, Frey & Embrechts (2005).

Transition Density of a Multi-dimensional Wiener Process

As an alternative to copulas one can express the dependence structure of the components of a stochastic process by its transition densities. Of course, for any given transition density there exists a corresponding copula. As a first example of a continuous multi-dimensional stochastic process, whose transition density can be expressed explicitly, we focus on the d-dimensional Wiener process, see Sect. 1.1. This fundamental stochastic process has a multivariate Gaussian transition density of the form

$$p(s, \boldsymbol{x}; t, \boldsymbol{y}) = \frac{1}{(2\pi(t-s))^{d/2}\sqrt{\det \boldsymbol{\Sigma}}} \exp\left\{\frac{(\boldsymbol{y}-\boldsymbol{x})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{x})}{2(t-s)}\right\}, \tag{2.2.9}$$

for $t \in [0, \infty)$, $s \in [0, t]$ and $\boldsymbol{x}, \boldsymbol{y} \in \Re^d$. Here $\boldsymbol{\Sigma}$ is a normalized covariance matrix. Its copula is the Gaussian copula (2.2.4), which is simply derived from the corresponding multi-variate Gaussian density. In the bivariate case with correlated Wiener processes this transition probability simplifies to

$$\begin{aligned} p(s, x_1, x_2; t, y_1, y_2) &= \frac{1}{2\pi(t-s)\sqrt{1-\varrho^2}} \\ &\times \exp\left\{-\frac{(y_1-x_1)^2 - 2(y_1-x_1)(y_2-x_2)\varrho + (y_2-x_2)^2}{2(t-s)\,(1-\varrho^2)}\right\}, \end{aligned} \tag{2.2.10}$$

for $t \in [0, \infty)$, $s \in [0, t]$ and $x_1, x_2, y_1, y_2 \in \Re$. Here the correlation parameter ϱ varies in the interval $[-1, 1]$. In the case of correlated Wiener processes one can first simulate independent Wiener processes and then form out of these, by linear transforms, correlated ones. Alternatively, one can follow the

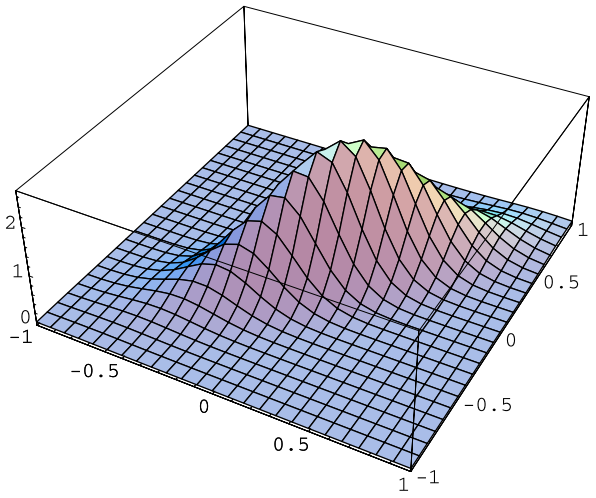

Fig. 2.2.1. Bivariate transition density of the two-dimensional Wiener process for fixed time step $\Delta = 0.1$, $x_1 = x_2 = 0.1$ and $\varrho = 0.8$

above inverse transform method by first using the Gaussian distribution for generating the increments of the first Wiener process component. Then one can condition the Gaussian distribution for the second component on these outcomes.

In Fig. 2.2.1 we illustrate the bivariate transition density of the two-dimensional Wiener process for the time increment $\Delta = t - s = 0.1$, initial values $x_1 = x_2 = 0.1$ and correlation $\varrho = 0.8$. One can also generate dependent Wiener processes that have a joint distribution with a given copula.

Transition Density of a Multi-dimensional Geometric Brownian Motion

The multi-dimensional geometric Brownian motion is a componentwise exponential of a linearly transformed Wiener process. Given a vector of correlated Wiener processes $\boldsymbol{W}$ with the transition density (2.2.9) we consider the following transformation

$$\boldsymbol{S}_t = \boldsymbol{S}_0 \exp\{\boldsymbol{a}t + \boldsymbol{B}\boldsymbol{W}_t\}, \tag{2.2.11}$$

for $t \in [0, \infty)$, where the exponential is taken componentwise. Here $\boldsymbol{a}$ is a vector of length d, while the elements of the matrix $\boldsymbol{B}$ are as follows

$$B^{i,j} = \begin{cases} b^j \text{ for } i = j \\ 0 \quad \text{otherwise,} \end{cases} \tag{2.2.12}$$

where $i, j \in \{1, 2, \ldots, d\}$.

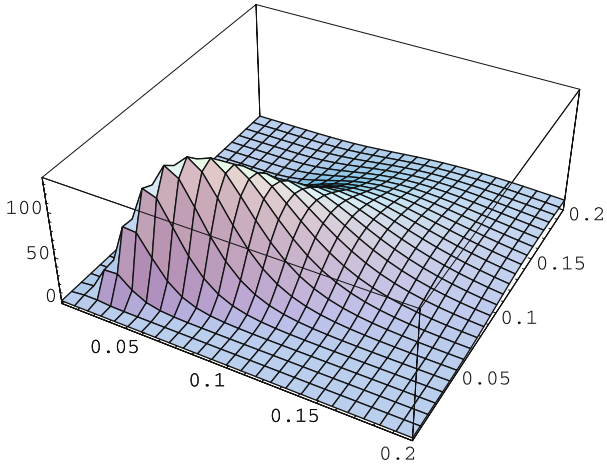

Fig. 2.2.2. Bivariate transition density of the two-dimensional geometric Brownian motion for $\Delta = 0.1$, $x_1 = x_2 = 0.1$, $\varrho = 0.8$, $b^1 = b^2 = 2$ and $a_1 = a_2 = 0.1$

Then the transition density of the above defined geometric Brownian motion has the following form

$$p(s,\boldsymbol{x};t,\boldsymbol{y}) = \frac{1}{(2\pi(t-s))^{d/2}\sqrt{\det \boldsymbol{\Sigma}}\prod_{i=1}^{d} b^i y_i} \tag{2.2.13}$$
$$\times \exp\left\{-\frac{(\ln(\boldsymbol{y}) - \ln(\boldsymbol{x}) - \boldsymbol{a}(t-s))^\top \boldsymbol{B}^{-1}\boldsymbol{\Sigma}^{-1}\boldsymbol{B}^{-1}(\ln(\boldsymbol{y}) - \ln(\boldsymbol{x}) - \boldsymbol{a}(t-s))}{2(t-s)}\right\}$$

for $t \in [0,\infty)$, $s \in [0,t]$ and $\boldsymbol{x}, \boldsymbol{y} \in \Re_+^d$. Here the logarithm is understood componentwise. In the bivariate case this transition density takes the particular form

$$\begin{aligned} p(s,x_1,x_2;t,y_1,y_2) &= \frac{1}{2\pi(t-s)\sqrt{1-\varrho^2}b^1b^2y_1y_2} \\ &\times \exp\left\{-\frac{(\ln(y_1) - \ln(x_1) - a^1(t-s))^2}{2(b^1)^2(t-s)(1-\varrho^2)}\right\} \\ &\times \exp\left\{-\frac{(\ln(y_2) - \ln(x_2) - a^2(t-s))^2}{2(b^2)^2(t-s)(1-\varrho^2)}\right\} \\ &\times \exp\left\{\frac{(\ln(y_1) - \ln(x_1) - a^1(t-s))(\ln(y_2) - \ln(x_2) - a^2(t-s))\varrho}{b^1b^2(t-s)(1-\varrho^2)}\right\}, \end{aligned}$$

for $t \in [0,\infty)$, $s \in [0,t]$ and $x_1, x_2, y_1, y_2 \in \Re^+$, where $\varrho \in [-1,1]$.

In Fig. 2.2.2 we illustrate the bivariate transition density of the two-dimensional geometric Brownian motion for the time increment $\Delta = t - s = 0.1$, initial values $x_1 = x_2 = 0.1$, correlation $\varrho = 0.8$, volatilities $b^1 = b^2 = 2$ and growth parameters $a^1 = a^2 = 0.1$.

Transition Density of a Multi-dimensional OU-Process

Another example is the standard d-dimensional *Ornstein-Uhlenbeck* (OU)-*process*. This process has a Gaussian transition density of the form

$$p(s,\boldsymbol{x};t,\boldsymbol{y}) = \frac{1}{(2\pi(1-e^{-2(t-s)}))^{d/2}\sqrt{\det \boldsymbol{\Sigma}}} \times \exp\left\{-\frac{(\boldsymbol{y}-\boldsymbol{x}e^{-(t-s)})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{x}e^{-(t-s)})}{2(1-e^{-2(t-s)})}\right\}, \quad (2.2.14)$$

for $t \in [0,\infty)$, $s \in [0,t]$ and $\boldsymbol{x},\boldsymbol{y} \in \Re^d$, with mean $\boldsymbol{x}e^{-(t-s)}$ and covariance matrix $\boldsymbol{\Sigma}(1-e^{-2(t-s)})$, $d \in \{1,2,\ldots\}$. In the bivariate case the transition density of the standard OU-process is expressed by

$$\begin{aligned} p(s,x_1,x_2;t,y_1,y_2) = {} & \frac{1}{2\pi\left(1-e^{-2(t-s)}\right)\sqrt{1-\varrho^2}} \\ & \times \exp\left\{-\frac{\left(y_1-x_1e^{-(t-s)}\right)^2+\left(y_2-x_2e^{-(t-s)}\right)^2}{2\left(1-e^{-2(t-s)}\right)(1-\varrho^2)}\right\} \\ & \times \exp\left\{\frac{\left(y_1-x_1e^{-(t-s)}\right)\left(y_2-x_2e^{-(t-s)}\right)\varrho}{\left(1-e^{-2(t-s)}\right)(1-\varrho^2)}\right\}, \end{aligned} \quad (2.2.15)$$

for $t \in [0,\infty)$, $s \in [0,t]$ and $x_1,x_2,y_1,y_2 \in \Re$, where $\varrho \in [-1,1]$.

Transition Density of a Multi-dimensional Geometric OU-Process

The transition density of a d-dimensional *geometric OU-process* can be obtained from the transition density of the multi-dimensional OU-process by applying the exponential transformation. Therefore, it can be expressed as

$$\begin{aligned} p(s,\boldsymbol{x};t,\boldsymbol{y}) = {} & \frac{1}{(2\pi(1-e^{-2(t-s)}))^{d/2}\sqrt{\det \boldsymbol{\Sigma}}\prod_{i=1}^d y_i} \\ & \times \exp\left\{-\frac{(\ln(\boldsymbol{y})-\ln(\boldsymbol{x})e^{-(t-s)})^\top \boldsymbol{\Sigma}^{-1}(\ln(\boldsymbol{y})-\ln(\boldsymbol{x})e^{-(t-s)})}{2(1-e^{-2(t-s)})}\right\}, \end{aligned} \quad (2.2.16)$$

for $t \in [0,\infty)$, $s \in [0,t]$ and $\boldsymbol{x},\boldsymbol{y} \in \Re_+^d$, $d \in \{1,2,\ldots\}$. In the bivariate case the transition density of the multi-dimensional OU-process is of the form

$$\begin{aligned} p(s,x_1,x_2;t,y_1,y_2) = {} & \frac{1}{2\pi\left(1-e^{-2(t-s)}\right)\sqrt{1-\varrho^2}y_1y_2} \\ & \times \exp\left\{-\frac{\left(\ln(y_1)-\ln(x_1)e^{-(t-s)}\right)^2+\left(\ln(y_2)-\ln(x_2)e^{-(t-s)}\right)^2}{2\left(1-e^{-2(t-s)}\right)(1-\varrho^2)}\right\} \\ & \times \exp\left\{\frac{\left(\ln(y_1)-\ln(x_1)e^{-(t-s)}\right)\left(\ln(y_2)-\ln(x_2)e^{-(t-s)}\right)\varrho}{\left(1-e^{-2(t-s)}\right)(1-\varrho^2)}\right\}, \end{aligned} \quad (2.2.17)$$

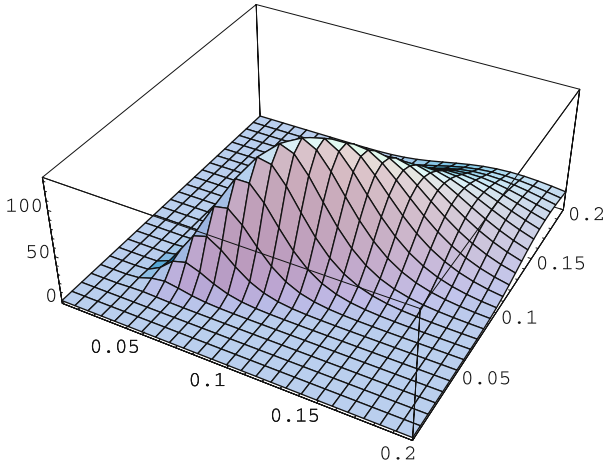

Fig. 2.2.3. Bivariate transition density of the two-dimensional geometric OU-process for $\Delta = 0.1$, $x_1 = x_2 = 0.1$ and $\varrho = 0.8$

for $t \in [0, \infty)$, $s \in [0, t]$ and $x_1, x_2, y_1, y_2 \in \Re_+$, where $\varrho \in [-1, 1]$.

In Fig. 2.2.3 we illustrate the bivariate transition density of the two-dimensional geometric OU-process for the time increment $\Delta = t - s = 0.1$, initial values $x_1 = x_2 = 0.1$ and correlation $\varrho = 0.8$. It is now obvious how to obtain the transition density of the componentwise exponential of other Gaussian vector processes.

Transition Density of a Wishart Process

Let us now study another class of processes that is related to products of Wiener processes. We characterize below the transition density of the matrix valued *Wishart process* for dimension parameter $\delta > 0$, starting at the time $s \in [0, \infty)$, in $\boldsymbol{X} > 0$ and being at time $t \in (s, \infty)$ in $\boldsymbol{Y}$, see Bru (1991) and Gouriéroux & Sufana (2004). Its transition density has the form

$$\begin{aligned} p_\delta(s, \boldsymbol{X}; t, \boldsymbol{Y}) &= \frac{1}{(2(t-s))^{\delta m/2} \boldsymbol{\Gamma}_m(\delta/2)} \operatorname{etr}\left\{-\frac{\boldsymbol{X}+\boldsymbol{Y}}{2(t-s)}\right\} (\det \boldsymbol{Y})^{(\delta-m-1)/2} \\ &\quad \times {}_0\boldsymbol{F}_1\left(\frac{\delta}{2}; \frac{\boldsymbol{X}\boldsymbol{Y}}{4(t-s)^2}\right) \\ &= \frac{1}{(2(t-s))^{m(m+1)/2}} \left(\frac{\det(\boldsymbol{Y})}{\det(\boldsymbol{X})}\right)^{\frac{\delta-m-1}{4}} \operatorname{etr}\left\{-\frac{\boldsymbol{X}+\boldsymbol{Y}}{2(t-s)}\right\} \\ &\quad \times \tilde{\boldsymbol{I}}_{(\delta-m-1)/2}\left(\frac{\boldsymbol{X}\boldsymbol{Y}}{4(t-s)^2}\right), \end{aligned} \tag{2.2.18}$$

see Donati-Martin, Doumerc, Matsumoto & Yor (2004). Here etr$\{\cdot\}$ denotes the elementwise exponential of the trace of a matrix and $\tilde{\boldsymbol{I}}_\nu$ is a special function of a matrix argument with index ν defined by

$$\tilde{\boldsymbol{I}}_\nu(\boldsymbol{z}) = \frac{(\det \boldsymbol{z})^{\nu/2}}{\boldsymbol{\Gamma}_m((m+1)/2+\nu)}\ {}_0\boldsymbol{F}_1((m+1)/2+\nu, \boldsymbol{z}). \tag{2.2.19}$$

It is related to the modified Bessel function of the first kind $\boldsymbol{I}_\nu(\cdot)$ by the relation $\tilde{\boldsymbol{I}}_\nu(\boldsymbol{z}) = \boldsymbol{I}_\nu(2\boldsymbol{z}^{1/2})$. In general, the hypergeometric function ${}_p\boldsymbol{F}_q$ in (2.2.19) can be expressed in terms of zonal polynomials, see Muirhead (1982). Finally, $\boldsymbol{\Gamma}_m(\cdot)$ denotes the multi-dimensional gamma function with

$$\boldsymbol{\Gamma}_m(x) = \int_{\Lambda>0} \text{etr}\{-\Lambda\}(\det(\Lambda))^{x-\frac{m+1}{2}} d\Lambda. \tag{2.2.20}$$

The transition density of the Wishart process, when it starts at time zero at $\boldsymbol{X} = \boldsymbol{0}$ for being at time $t \in (0,\infty)$ in $\boldsymbol{Y} \geq \boldsymbol{0}$, can be written as

$$p_\delta(0, \boldsymbol{0}; t, \boldsymbol{Y}) = (2\,t)^{-\delta m/2}\, \frac{\det(\boldsymbol{Y})^{(\delta-m-1)/2}}{\boldsymbol{\Gamma}_m(\frac{\delta}{2})}\, \text{etr}\left\{-\frac{\boldsymbol{Y}}{2\,t}\right\}. \tag{2.2.21}$$

Herz (1955) derived a representation of the non-central Wishart density in terms of a Bessel function of matrix argument $A_\nu^{(m)}$. The advantage of this representation is that it also accounts for correlation. It has the form

$$p_\delta(s, \boldsymbol{X}; t, \boldsymbol{Y}) = \frac{1}{(2(t-s)\det(\boldsymbol{\Sigma}))^{\delta/2}} \tag{2.2.22}$$
$$\times\, \text{etr}\left\{-\frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}+\boldsymbol{Y})}{2(t-s)}\right\} \det(\boldsymbol{Y})^{(\delta-m-1)/2} A^{(m)}_{(\delta-m-1)/2}\left(-\frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}}{4(t-s)^2}\right),$$

where $\boldsymbol{\Sigma}$ is a normalized covariance matrix. Herz (1955) provided also the following representation for the special function

$$A^{(2)}_\nu(\boldsymbol{z}) = \frac{1}{\sqrt{\pi}} \sum_{j=0}^{\infty} \frac{1}{j!\,\Gamma(\nu+j+1)}\ A^{(1)}_{\nu+2j+1/2}(\text{tr}(\boldsymbol{z}))\det(\boldsymbol{z})^j, \tag{2.2.23}$$

where $A^{(1)}_\nu(z) = \sum_{j=0}^{\infty}(-z)^j/(j!\,\Gamma(\nu+j+1))$. Here $\Gamma(\cdot)$ denotes the well-known gamma function.

Hence, with the use of (2.2.22) and (2.2.23) we obtain for $m = 2$ the following transition density for the 2×2 Wishart process of dimension $\delta = 2$:

$$p(s, x_{11}, x_{12}, x_{21}, x_{22}; t, y_{11}, y_{12}, y_{21}, y_{22}) = \frac{{}_0F_3\left(\frac{1}{3}, \frac{2}{3}, 1; K\right)}{2\pi(s-t)^2(1-\varrho^2)\sqrt{y_{11}y_{22}-y_{12}y_{21}}}$$
$$\times \exp\left\{-\frac{x_{11}-\varrho x_{12}-\varrho x_{21}+x_{22}+y_{11}-\varrho y_{12}-\varrho y_{21}+y_{22}}{2(s-t)\,(\varrho^2-1)}\right\}$$
$$\times \cosh\left(\frac{1}{2}\sqrt{\frac{(x_{12}x_{21}-x_{11}x_{22})\,(y_{12}y_{21}-y_{11}y_{22})}{(s-t)^4\,(\varrho^2-1)^2}}\right), \tag{2.2.24}$$

where

$$K = \frac{1}{108(s-t)^2(\varrho^2-1)^2}(x_{22}y_{11}\varrho^2 + x_{12}y_{12}\varrho^2 - x_{12}y_{11}\varrho - x_{22}y_{12}\varrho - x_{22}y_{21}\varrho - x_{12}y_{22}\varrho + x_{12}y_{21} + x_{21}(y_{12} + \varrho(-y_{11} + \varrho y_{21} - y_{22})) + x_{22}y_{22} + x_{11}(y_{11} + \varrho(-y_{12} - y_{21} + \varrho y_{22})))$$

and ${}_pF_q(\cdot,\cdot,\cdot;K)$ is a generalized hypergeometric function explained in Muirhead (1982).

Transition Density of a Square Root Process

Similarly one can characterize the transition density for a *square root* (SR)-*process*, see (2.1.3).

A scalar square root (SR)-process $X = \{X_t, t \geq 0\}$ of dimension $\delta > 2$ is given by the SDE

$$dX_t = \left(\frac{\delta}{4}c_t^2 + b_t X_t\right)dt + c_t\sqrt{X_t}\,dW_t \tag{2.2.25}$$

for $t \geq 0$ with $X_0 > 0$. Here $W = \{W_t, t \geq 0\}$ is a scalar Wiener process and c_t and b_t are deterministic functions of time. Let

$$s_t = \exp\left\{\int_0^t b_u\,du\right\} \tag{2.2.26}$$

and

$$\varphi_t = \frac{1}{4}\int_0^t \frac{c_u^2}{s_u}\,du \tag{2.2.27}$$

for $t \geq 0$. Then the transition density of the scalar SR-process X is given in the form

$$p(s,x;t,y) = \frac{1}{2\,s_t\,\varphi_t}\left(\frac{y}{x\,s_t}\right)^{\frac{\nu}{2}}\exp\left\{-\frac{x+\frac{y}{s_t}}{2\,\varphi_t}\right\}I_\nu\left(\frac{\sqrt{x\,\frac{y}{s_t}}}{\varphi_t}\right) \tag{2.2.28}$$

for $0 \leq s < t < \infty$ and $x, y \in (0,\infty)$, where $I_\nu(\cdot)$ is the modified Bessel function of the first kind with index $\nu = \frac{\delta}{2} - 1$.

Transition Density of a Matrix SR-process

For the multi-dimensional SR-process, given in terms of a $d \times m$ matrix, we have an analytic transition density that can be derived from (2.2.18) or (2.2.22) in the form

$$p(s,\boldsymbol{X};t,\boldsymbol{Y}) = \frac{p_\delta\left(\varphi(s),\frac{\boldsymbol{X}}{s_s};\varphi(t),\frac{\boldsymbol{Y}}{s_t}\right)}{s_t} \tag{2.2.29}$$

for $0 \le s < t < \infty$ and nonnegative elements in the $d \times m$ matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$. Here we use the transformed φ-time with

$$\varphi(t) = \varphi(0) + \frac{\bar{b}^2}{4\bar{c}s_0}\left(1 - \exp\{-\bar{c}t\}\right) \tag{2.2.30}$$

and $s_t = s_0 \exp\{\bar{c}t\}$ for $t \in [0, \infty)$, $s_0 > 0$, $\bar{c} < 0$ and $\bar{b} \neq 0$.

As an example, we write down the transition density of the 2×2 matrix SR-process of dimension $\delta = 2$. This density follows from (2.2.24) and it can be expressed by

$$\begin{aligned}
&p(s, x_{11}, x_{12}, x_{21}, x_{22}; t, y_{11}, y_{12}, y_{21}, y_{22}) \\
&= \frac{8\bar{c}^2 s_0^2 e^{\bar{c}(2s+t)}\, {}_0F_3\left(\frac{1}{3}, \frac{2}{3}, 1; K\right)}{\bar{b}^4 \left(e^{\bar{c}s} - e^{\bar{c}t}\right)^2 \pi \left(1 - \varrho^2\right) \sqrt{e^{-2\bar{c}t}\left(y_{11}y_{22} - y_{12}y_{21}\right)}} \\
&\times \exp\left\{\frac{-2\bar{c}e^{\bar{c}t}\left(x_{11} - \varrho\left(x_{12} + x_{21}\right) + x_{22}\right) - 2\bar{c}e^{\bar{c}s}\left(y_{11} - \varrho\left(y_{12} + y_{21}\right) + y_{22}\right)}{\bar{b}^2\left(e^{\bar{c}s} - e^{\bar{c}t}\right)\left(\varrho^2 - 1\right)}\right\} \\
&\times \cosh\left(8\sqrt{\frac{\bar{c}^4 e^{2\bar{c}(s+t)}\left(x_{12}x_{21} - x_{11}x_{22}\right)\left(y_{12}y_{21} - y_{11}y_{22}\right)}{\bar{b}^8\left(e^{\bar{c}s} - e^{\bar{c}t}\right)^4\left(\varrho^2 - 1\right)^2}}\right),
\end{aligned} \tag{2.2.31}$$

where

$$\begin{aligned}
K = \frac{4\bar{c}^2 e^{\bar{c}(s+t)}}{27\bar{b}^4\left(e^{\bar{c}s} - e^{\bar{c}t}\right)^2\left(\varrho^2 - 1\right)^2}\big(&x_{22}y_{11}\varrho^2 + x_{21}y_{21}\varrho^2 - x_{21}y_{11}\varrho - x_{22}y_{12}\varrho \\
&- x_{22}y_{21}\varrho - x_{21}y_{22}\varrho + x_{21}y_{12} + x_{12}\left(y_{21} + \varrho\left(-y_{11} + \varrho y_{12} - y_{22}\right)\right) \\
&+ x_{22}y_{22} + x_{11}\left(y_{22}\varrho^2 - \left(y_{12} + y_{21}\right)\varrho + y_{11}\right)\big).
\end{aligned}$$

The above transition density looks complex. However, it has the great advantage of providing an explicit formula, which is extremely valuable in many quantitative investigations.

Transition Density of Multi-dimensional Lévy Processes

Lévy processes are a special class of processes which have independent and stationary increments, see Sect. 1.7. It turns out that the distributions of the independent increments of Lévy processes are infinitely divisible, see Cont & Tankov (2004). A widely used class of distributions for the increments with this property are members of the family of generalized hyperbolic (GH) distributions. It is always possible to construct a Lévy process so that the value of the increment of the process over a fixed time interval has a given GH distribution. The GH density can be expressed as

$$p(\boldsymbol{x}) = c\frac{K_{\lambda-(\frac{d}{2})}\left(\sqrt{(\chi+(\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))(\psi+\boldsymbol{\gamma}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)e^{(\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma}}}{\left(\sqrt{(\chi+(\boldsymbol{x}-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}))(\psi+\boldsymbol{\gamma}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})}\right)^{(\frac{d}{2})-\lambda}}, \tag{2.2.32}$$

where

$$c = \frac{(\chi\psi)^{-\lambda}\,\psi^\lambda(\psi+\boldsymbol{\gamma}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\gamma})^{(d/2)-\lambda}}{(2\pi)^{d/2}\det(\boldsymbol{\Sigma})^{1/2}K_\lambda\left(\sqrt{\chi\psi}\right)}, \tag{2.2.33}$$

$\boldsymbol{\Sigma}$ is a correlation type matrix, $\boldsymbol{\mu}$ a drift vector, $\boldsymbol{\gamma}$ a scaling vector, and χ, ψ and λ are shape parameters.

$K_\lambda(\cdot)$ denotes a modified Bessel function of the third kind and $\boldsymbol{x} \in \Re^d$. Additionally, if $\boldsymbol{\gamma} = \boldsymbol{0}$, then this distribution is symmetric.

Some known special cases of this distribution include the d-dimensional hyperbolic distribution if $\lambda = \frac{1}{2}(d+1)$; the d-dimensional normal inverse Gaussian (NIG) distribution if $\lambda = -\frac{1}{2}$; the d-dimensional variance-gamma (VG) distribution if $\lambda > 0$ and $\chi = 0$ and the d-dimensional skewed Student-t distribution if $\lambda = -\frac{1}{2}\nu$, $\chi = \nu$ and $\psi = 0$. Note that the last two special cases are limiting distributions.

In the two-dimensional case the GH density simplifies to

$$p(x_1, x_2) = c\,\exp\{A\}\,\frac{K_{\lambda-1}(\xi)}{\xi^{1-\lambda}}, \tag{2.2.34}$$

where

$$A = \frac{\gamma_1\mu_1 - \gamma_2\varrho\mu_1 + \gamma_2\mu_2 - \gamma_1\mu_2\varrho - \gamma_1 x_1 + \gamma_2\varrho x_1 - \gamma_2 x_2 + \gamma_1\varrho\, x_2}{\varrho^2 - 1},$$

$$\xi = \sqrt{(-\gamma_1^2 + 2\gamma_2\varrho\gamma_1 - \gamma_2^2 + \varrho^2\psi - \psi)\,B}\,,$$

$$B = \frac{(\mu_1^2 - 2\mu_2\varrho\mu_1 + \mu_2^2 - x_1^2 - x_2^2 + \varrho^2\chi - \chi + 2\varrho\, x_1 x_2)}{(\varrho^2-1)^2}$$

and

$$c = \frac{\left(\sqrt{\chi\psi}\right)^{-\lambda}\psi^\lambda\left(\frac{\gamma_1^2 - 2\gamma_2\varrho\gamma_1 + \gamma_2^2 - \varrho^2\psi + \psi}{1-\varrho^2}\right)^{1-\lambda}}{2\pi\sqrt{1-\varrho^2}K_\lambda\left(\sqrt{\chi\psi}\right)}. \tag{2.2.35}$$

In Fig. 2.2.4 we illustrate a bivariate GH density with $\mu_1 = \mu_2 = 0.1$, $\gamma_1 = \gamma_2 = 0.2$, $\chi = \psi = 0.4$, $\lambda = -0.5$ and $\varrho = 0.8$. One notes the much fatter tails of the density in Fig. 2.2.4 when compared with those of the Wiener process increment shown in Fig. 2.2.1. Extreme joint events are far more likely. In principle, one can generate paths of multi-dimensional Lévy processes for a given copula and given marginal transition densities as discussed earlier.

Since we do not discuss the simulation of Lévy processes including those with infinite intensity with great detail, we refer the reader for details on this

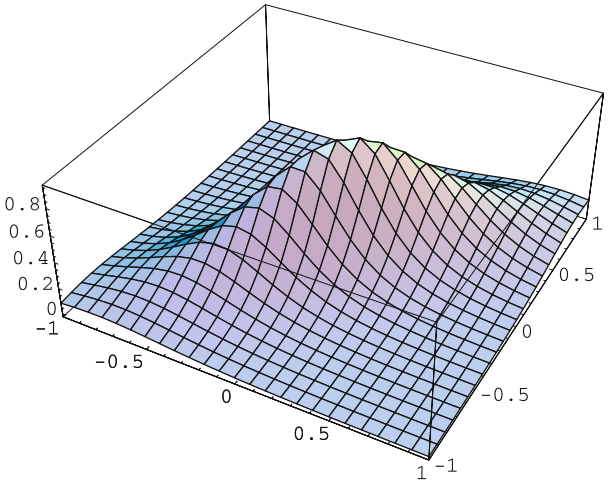

Fig. 2.2.4. Bivariate GH density for $\mu_1 = \mu_2 = 0.1$, $\gamma_1 = \gamma_2 = 0.2$, $\chi = \psi = 0.4$, $\lambda = -0.5$ and $\varrho = 0.8$

subject to Madan & Seneta (1990), Eberlein & Keller (1995), Bertoin (1996), Protter & Talay (1997), Barndorff-Nielsen & Shephard (2001), Geman et al. (2001), Eberlein (2002), Kou (2002), Rubenthaler (2003), Cont & Tankov (2004), Jacod, Kurtz, Méléard & Protter (2005) and Klüppelberg, Lindner & Maller (2006).

Other Explicit Transition Densities

There is still a wider range of one-dimensional Markov processes with explicit transition densities than those mentioned so far. For instance, Craddock & Platen (2004) consider generalized square root processes, see also Platen & Heath (2006). These are diffusion processes $X = \{X_t, t \in [0, \infty)\}$ with a square root function $b(t, x) = \sqrt{2x}$ as diffusion coefficient for all $t \geq 0$ and $x \in [0, \infty)$. By exploiting Lie group symmetries they identified a collection of drift functions $a(t, x) = a(x)$ for which one still has an analytic formula for the corresponding transition density $p(0, x; t, y)$. These include the cases of the following drifts yielding corresponding transition densities that we list below:

(i)

$$a(x) = \alpha > 0,$$

$$p(0, x; t, y) = \frac{1}{t} \left(\frac{x}{y} \right)^{\frac{1-\alpha}{2}} I_{\alpha-1} \left(\frac{2\sqrt{x\,y}}{t} \right) \exp\left\{ -\frac{(x+y)}{t} \right\};$$

(ii)

$$a(x) = \frac{\mu x}{1 + \frac{\mu}{2} x},\ \mu > 0,$$

$$p(0,x;t,y) = \frac{\exp\left\{-\frac{(x+y)}{t}\right\}}{\left(1+\frac{\mu}{2}\,x\right)t}\left[\left(\sqrt{\frac{x}{y}} + \frac{\mu\sqrt{x\,y}}{2}\right) I_1\left(\frac{2\sqrt{x\,y}}{t}\right) + t\,\delta(y)\right];$$

(iii)

$$a(x) = \frac{1+3\sqrt{x}}{2(1+\sqrt{x})},$$

$$p(0,x;t,y) = \frac{\cosh\left(\frac{2\sqrt{x\,y}}{t}\right)}{\sqrt{\pi\,y\,t}\,(1+\sqrt{x})}\left(1+\sqrt{y}\,\tanh\left(\frac{2\sqrt{x\,y}}{t}\right)\right)$$
$$\times\ \exp\left\{-\frac{(x+y)}{t}\right\};$$

(iv)

$$a(x) = 1 + \mu\tanh\left(\mu + \frac{1}{2}\mu\ln(x)\right),\ \mu = \frac{1}{2}\sqrt{\frac{5}{2}},$$

$$p(0,x;t,y) = \left(\frac{x}{y}\right)^{\frac{\mu}{2}}\left[I_{-\mu}\left(\frac{2\sqrt{x\,y}}{t}\right) + e^{2\,\mu}\,y^{\mu}\,I_{\mu}\left(\frac{2\sqrt{x\,y}}{t}\right)\right]$$
$$\times\ \frac{\exp\{-\frac{x+y}{t}\}}{(1+\exp\{2\,\mu\}\,x^{\mu})\,t};$$

(v)

$$a(x) = \frac{1}{2} + \sqrt{x},$$

$$p(0,x;t,y) = \cosh\left(\frac{(t+2\sqrt{x})\sqrt{y}}{t}\right)\frac{\exp\{-\sqrt{x}\}}{\sqrt{\pi\,y\,t}}\,\exp\left\{-\frac{(x+y)}{t} - \frac{t}{4}\right\};$$

(vi)

$$a(x) = \frac{1}{2} + \sqrt{x}\tanh(\sqrt{x}),$$

$$p(0,x;t,y) = \frac{\cosh\left(\frac{2\sqrt{x\,y}}{t}\right)}{\sqrt{\pi\,y\,t}}\,\frac{\cosh(\sqrt{y})}{\cosh(\sqrt{x})}\,\exp\left\{-\frac{(x+y)}{t} - \frac{t}{4}\right\};$$

(vii)

$$a(x) = \frac{1}{2} + \sqrt{x}\coth(\sqrt{x}),$$

$$p(0,x;t,y) = \frac{\sinh\left(\frac{2\sqrt{x\,y}}{t}\right)}{\sqrt{\pi\,y\,t}}\,\frac{\sinh(\sqrt{y})}{\sinh(\sqrt{x})}\,\exp\left\{-\frac{(x+y)}{t} - \frac{t}{4}\right\};$$

(viii)

$$a(x) = 1 + \cot(\ln(\sqrt{x})) \quad \text{for } x \in (\exp\{-2\pi\}, 1),$$

$$p(0,x;t,y) = \frac{\exp\{-\frac{(x+y)}{t}\}}{2\,\imath\, t \sin(\ln(\sqrt{x}))} \left(y^{\frac{\imath}{2}}\, I_{\imath}\left(\frac{2\sqrt{x\,y}}{t}\right) - y^{-\frac{\imath}{2}}\, I_{-\imath}\left(\frac{2\sqrt{x\,y}}{t}\right)\right);$$

(ix)

$$a(x) = x \coth\left(\frac{x}{2}\right),$$

$$p(0,x;t,y) = \frac{\sinh(\frac{y}{2})}{\sinh(\frac{x}{2})} \exp\left\{-\frac{(x+y)}{2\,\tanh(\frac{t}{2})}\right\} \times \left[\frac{\exp\{\frac{t}{2}\}}{\exp\{t\}-1}\sqrt{\frac{x}{y}}\, I_1\left(\frac{\sqrt{x\,y}}{\sinh(\frac{t}{2})}\right) + \delta(y)\right];$$

(x)

$$a(x) = x \tanh\left(\frac{x}{2}\right),$$

$$p(0,x;t,y) = \frac{\cosh(\frac{y}{2})}{\cosh(\frac{x}{2})} \exp\left\{-\frac{(x+y)}{2\,\tanh(\frac{t}{2})}\right\} \times \left[\frac{\exp\{\frac{t}{2}\}}{\exp\{t\}-1}\sqrt{\frac{x}{y}}\, I_1\left(\frac{\sqrt{x\,y}}{\sinh(\frac{t}{2})}\right) + \delta(y)\right].$$

Here I_α is the modified Bessel function of the first kind with index α, see Platen & Heath (2006) and (2.2.19).

The first drift above in (i) is that of the well-known squared Bessel process. However, the other drifts are mostly from previously unknown scalar diffusion processes. It is obvious that some multi-dimensional versions of most of these processes can be constructed. It is an area of ongoing research that identifies natural dependence structures between different components of diffusions with explicit transition densities. One can, of course, always start from a given copula and generate dependent diffusions with given marginal transition distributions. An easy case is obtained, when all components of the multi-dimensional diffusion are independent. In general, each component of a multi-dimensional diffusion can come from various types of conditional transition distribution functions. This gives a great variety of multi-dimensional models with paths that can be exactly simulated via the inverse transform method.

Illustration of the Inverse Transform Method

Let us illustrate simulation via the inverse transform method for the following simple example of a two-dimensional standard OU-process. The marginal

transition distribution function of this process for its first component can be represented as

$$
\begin{aligned}
F_{Y_1}(s,x_1,t,y_1) &= \frac{1}{2}\left(1-e^{-2(t-s)}\right)\sqrt{1-\varrho^2} \\
&\quad\times\left[1-\operatorname{erf}\left\{x_1 e^{-(t-s)}\sqrt{\frac{1}{2}\left(1-e^{-2(t-s)}\right)}\right\}\right. \\
&\quad+ e^{-(t-s)}\sqrt{1-e^{-2(t-s)}}\left(x_1\sqrt{\frac{e^{2(t-s)}-1}{x_1^2}}\operatorname{erf}\left\{\sqrt{\frac{x_1^2}{2(e^{2(t-s)}-1)}}\right\}\right. \\
&\quad\left.\left.+\left(y_1e^{t-s}-x_1\right)\sqrt{\frac{1-e^{-2(t-s)}}{x_1e^{-(t-s)}-y_1}}\operatorname{erf}\left\{\frac{x_1e^{-(t-s)}-y_1}{2\left(1-e^{-2(t-s)}\right)}\right\}\right)\right],
\end{aligned}
\tag{2.2.36}
$$

for $x_1, y_1 \in \Re$ and $t > s$. Recall that $\operatorname{erf}\{\cdot\}$ denotes the well-known error function, see Abramowitz & Stegun (1972). Additionally, the conditional transition distribution function of the second component given the first component is

$$
\begin{aligned}
F_{Y_2|Y_1}(s,x_1,x_2,t,y_1,y_2) &= \frac{1}{2}\sqrt{\frac{1}{(1-e^{-2(t-s)})(1-\varrho^2)}} \\
&\times\left[\sqrt{(1-e^{-2(t-s)})(1-\varrho^2)}-\left(x_2-x_1\varrho-e^{t-s}(y_2-y_1\varrho)\right)\right. \\
&\left.\times\sqrt{\frac{(1-e^{-2(t-s)})(1-\varrho^2)}{\left(x_2-x_1\varrho-e^{t-s}(y_2-y_1\varrho)\right)^2}}\operatorname{erf}\left\{\sqrt{\frac{\left(x_2-x_1\varrho-e^{t-s}(y_2-y_1\varrho)\right)^2}{2(e^{2(t-s)}-1)(1-\varrho^2)}}\right\}\right],
\end{aligned}
\tag{2.2.37}
$$

for $x_1, x_2, y_1, y_2 \in \Re$, $\varrho \in [-1,1]$ and $t > s$. In Fig. 2.2.5 we display the marginal transition distribution function (2.2.36) for fixed $x_1 = 0.1$, while in Fig. 2.2.6 we illustrate the conditional transition distribution function (2.2.37) for fixed initial values $x_1 = x_2 = 0.1$ and the time increment $\Delta = t-s = 0.1$. In both graphs we assume the correlation $\varrho = 0.8$. We use the inverse transform method for the simulation of the two-dimensional OU-process and display a resulting trajectory of the related two components in Fig. 2.2.7. This example illustrates that we need a collection of marginal and conditional transition distribution functions when generating the components of the process one after the other. In the above Gaussian example this involved the use of the $\operatorname{erf}\{\cdot\}$ special function. For squared Bessel processes the non-central chi-square distribution function represents the analogue special function that needs to be employed. Otherwise, the procedure for simulating exact paths is very similar.

2.3 Exact Solutions of Multi-dimensional SDEs

Sometimes, for multi-dimensional distribution functions, as those introduced in Sect. 2.2, it may be more convenient to sample from the known multi-

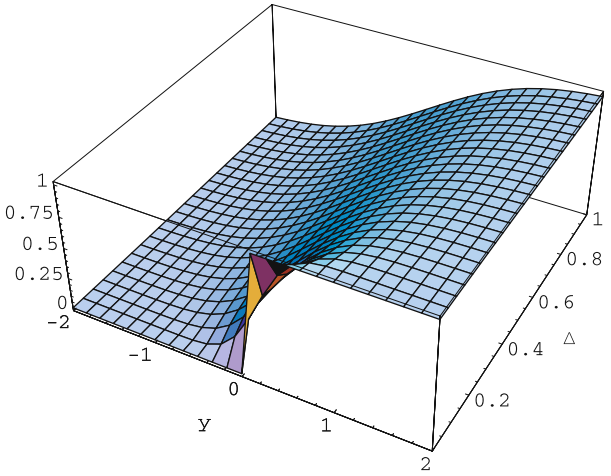

Fig. 2.2.5. Marginal transition distribution function of the first component y_1 of the two-dimensional OU-process for fixed $x_1 = 0.1$ in dependence on the time step Δ

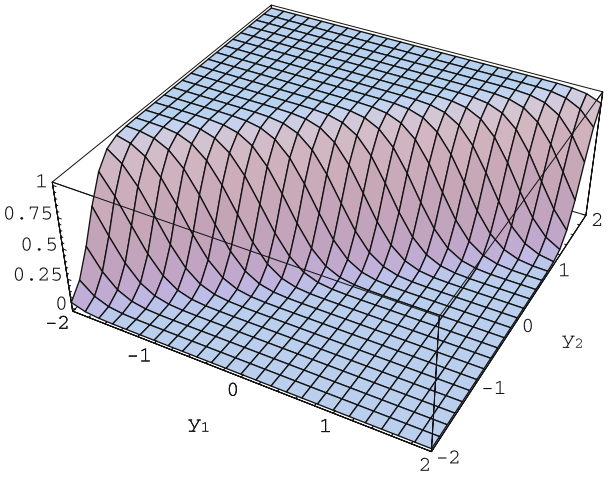

Fig. 2.2.6. Conditional transition distribution function of the second component y_2 of the two-dimensional OU-process given the first component y_1 for fixed $x_1 = x_2 = 0.1$ and $\Delta = 0.1$

dimensional distribution function directly rather than using the inverse-transform method. For instance, the increment of the multi-dimensional Wiener process can be simulated from the following exact relation

$$\boldsymbol{X}_{t+\Delta} - \boldsymbol{X}_t \sim N_d(\boldsymbol{0}, \boldsymbol{\Sigma}\Delta), \tag{2.3.1}$$

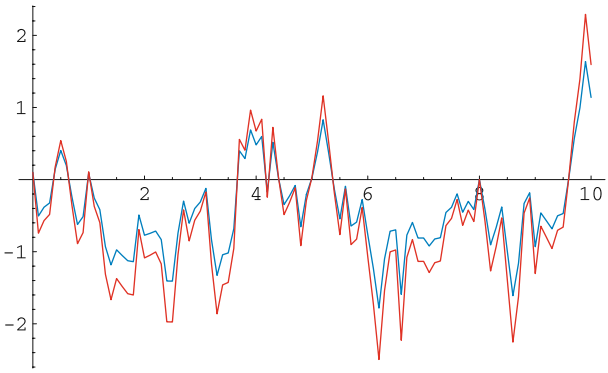

Fig. 2.2.7. Trajectory of the two components of the two-dimensional OU-process for $\Delta = 0.1$, initial value $\boldsymbol{x} = (0.1, 0.1)$ and $\varrho = 0.8$

where N_d denotes a d-dimensional Gaussian distribution with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}\Delta$. Similarly, the value at time $t + \Delta$ of the standard d-dimensional OU-process can be obtained by the relation

$$\boldsymbol{X}_{t+\Delta} \sim N_d(\boldsymbol{X}_t e^{-\Delta}, \boldsymbol{\Sigma}(1 - e^{-2\Delta})). \tag{2.3.2}$$

The $m \times m$ Wishart process can be simulated from the non-central Wishart distribution W_m with δ degrees of freedom, covariance matrix $\boldsymbol{\Sigma}\Delta$ and non-centrality matrix $\boldsymbol{\Sigma}^{-1}\boldsymbol{X}_t\Delta^{-1}$, where

$$\boldsymbol{X}_{t+\Delta} \sim W_m(\delta, \boldsymbol{\Sigma}\Delta, \boldsymbol{\Sigma}^{-1}\boldsymbol{X}_t\Delta^{-1}). \tag{2.3.3}$$

For details on how to sample conveniently from the non-central Wishart distribution we refer to Gleser (1976).

The increments of a Lévy processes constructed from a GH distribution can be obtained by

$$\boldsymbol{X}_{t+\Delta} - \boldsymbol{X}_t \sim GH_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma) \tag{2.3.4}$$

for fixed $\Delta > 0$. GH random variables can be simulated by subordination from d-dimensional Gaussian random variables whose mean and covariance matrix are made random in an appropriate way. This yields a mixture of normal increments by the relation

$$\boldsymbol{X}_{t+\Delta} - \boldsymbol{X}_t \sim \boldsymbol{\mu} + W\boldsymbol{\gamma} + \sqrt{W}\boldsymbol{A}\boldsymbol{Z}. \tag{2.3.5}$$

Here $\boldsymbol{Z} \sim N_k(\mathbf{0}, \mathbf{I}_k)$, where $W \geq 0$ is a non-negative scalar generalized inverse Gaussian (GIG) random variable that is independent of $\boldsymbol{Z}$. Here $\boldsymbol{I}_k$ denotes a k-dimensional unit matrix. $\boldsymbol{A}$ is a $d \times k$ matrix and $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$ are d-dimensional vectors. Hence, the conditional distribution of increment $\boldsymbol{X}_{t+\Delta} - \boldsymbol{X}_t$ given $W = w$ is conditionally Gaussian $N_d(\boldsymbol{\mu} + W\boldsymbol{\gamma}, w\boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^\top$.

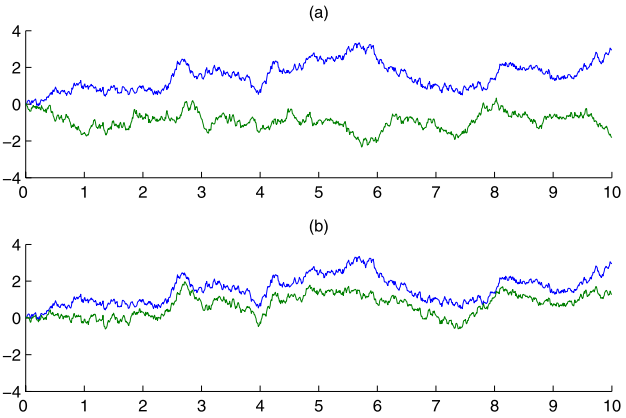

Fig. 2.3.1. (a) The trajectory of the independent components of a two-dimensional standard Wiener process; (b) Trajectories of two correlated Wiener processes

Let us now consider some selected multi-dimensional stochastic processes commonly used for modeling in finance but also popular in other areas of application. Our choice of these processes stems from the fact that they can all be simulated exactly, at least for some special cases.

Wiener Processes

The most important multi-dimensional continuous process with stationary independent increments is the d-dimensional standard Wiener process, see Sect. 1.1. It is a continuous process with independent Gaussian increments. First we assume that the components of the d-dimensional Wiener process $\boldsymbol{W}$, are independent. The increments of the Wiener processes $W_t^j - W_s^j$ for $j \in \{1, 2, \ldots, d\}$, $t \geq 0$ and $s \leq t$ are then independent Gaussian random variables with mean zero and variance equal to $t - s$. Therefore, one obtains the vector increments of the standard d-dimensional Wiener process $\boldsymbol{W}_t - \boldsymbol{W}_s \sim N_d(\boldsymbol{0}, (t-s)\boldsymbol{I})$ as a vector of zero mean independent Gaussian random variables with variance $t - s$. $\boldsymbol{I}$ denotes here the unit matrix. For the values of the trajectory of the standard d-dimensional Wiener process at the discretization times $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$, with $\Delta > 0$ we obtain the following iterative formula

$$\begin{aligned} \boldsymbol{W}_0 &= \boldsymbol{0} \\ \boldsymbol{W}_{t_{i+1}} &= \boldsymbol{W}_{t_i} + \sqrt{\Delta}\boldsymbol{N}_{i+1}, \end{aligned} \tag{2.3.6}$$

where $\boldsymbol{N}_{i+1} \sim N_d(\boldsymbol{0}, \boldsymbol{I})$ is an independent standard Gaussian random vector and $\boldsymbol{0}$ denotes the corresponding vector of zeros. We display in Fig. 2.3.1 (a) the trajectories of the independent components of a two-dimensional standard Wiener process.

Correlated Wiener Processes

Let us now define a d-dimensional continuous process

$$\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_t = (\tilde{W}_t^1, \tilde{W}_t^2, \ldots, \tilde{W}_t^d)^\top, t \in [0, \infty)\} \tag{2.3.7}$$

such that its components $\tilde{W}_t^1, \tilde{W}_t^2, \ldots, \tilde{W}_t^d$ are transformed scalar Wiener processes. In vector notation such a d-dimensional transformed Wiener process can be expressed by the linear transform

$$\tilde{\boldsymbol{W}}_t = \boldsymbol{a}t + \boldsymbol{B}\boldsymbol{W}_t, \tag{2.3.8}$$

where $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^\top$ is a d-dimensional vector and $\boldsymbol{B}$ is a $d \times m$-matrix and $\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^m)^\top$ is an m-dimensional standard Wiener process. Note that the kth component of $\tilde{\boldsymbol{W}}$ is such that

$$\tilde{W}_t^k = a_k t + \sum_{i=1}^{m} b_{k,i} W_t^i, \tag{2.3.9}$$

for $k \in \{1, 2, \ldots, d\}$. This means that $\tilde{W}_t^k$, $k \in \{1, 2, \ldots, d\}$, is constructed as a linear combination of components of the vector $\boldsymbol{W}_t$ plus some trend. From the properties of Gaussian random variables, the following relation emerges

$$\begin{aligned} \tilde{\boldsymbol{W}}_0 &= \boldsymbol{0}, \\ \tilde{\boldsymbol{W}}_{t_{i+1}} &= \tilde{\boldsymbol{W}}_{t_i} + \boldsymbol{a}\Delta + \sqrt{\Delta}\,\tilde{\boldsymbol{N}}_{i+1}, \end{aligned} \tag{2.3.10}$$

for $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$ with $\Delta > 0$. Here the random vector $\tilde{\boldsymbol{N}}_{i+1} \sim N_d(\boldsymbol{0}, \boldsymbol{\Sigma})$ is a d-dimensional Gaussian vector with $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$, which is independent for each $i \in \{0, 1, \ldots\}$.

We display in Fig. 2.3.1 (b) a trajectory of a two-dimensional transformed Wiener process $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_t = (\tilde{W}_t^1, \tilde{W}_t^2)^\top, t \in [0, 10]\}$ with correlated components. The two components of this process are as follows

$$\tilde{W}_t^1 = W_t^1, \tag{2.3.11}$$

$$\tilde{W}_t^2 = \varrho W_t^1 + \sqrt{1 - \varrho^2} W_t^2, \tag{2.3.12}$$

for $t \in [0, 10]$ with correlation $\varrho = 0.8$. In Fig. 2.3.1 (a) we showed the corresponding independent Wiener paths W_t^1 and W_t^2 for $t \in [0, 10]$. We note the expected strong similarity between $\tilde{W}_t^1$ and $\tilde{W}_t^2$ in Fig. 2.3.1 (b). The two-dimensional Wiener process $\tilde{\boldsymbol{W}}$ can also be expressed using matrix multiplication as $\tilde{\boldsymbol{W}}_t = \boldsymbol{B}\boldsymbol{W}_t$, where

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ \varrho & \sqrt{1 - \varrho^2} \end{pmatrix} \tag{2.3.13}$$

and $\boldsymbol{W}_t = (W_t^1, W_t^2)^\top$.

Matrix Wiener Processes

As we observed already with Wishart processes, matrix valued processes may be convenient in a given context. Therefore, let us define a $d \times m$ standard matrix Wiener process $\boldsymbol{W} = \{\boldsymbol{W}_t = [W_t^{i,j}]_{i,j=1}^{d,m}, t \in [0,\infty)\}$. This matrix stochastic process can be obtained by the following construction

$$\begin{aligned} \boldsymbol{W}_0 &= \boldsymbol{0} \\ \boldsymbol{W}_{t_{i+1}} &= \boldsymbol{W}_{t_i} + \sqrt{\Delta}\boldsymbol{N}_{i+1}, \end{aligned} \tag{2.3.14}$$

for $t_i = i\Delta$, $i = \{0, 1, \ldots\}$ with $\Delta > 0$ and $d \times m$-matrix $\boldsymbol{0}$ of zero elements. Here $\boldsymbol{N}_{i+1} \sim N_{d\times m}(\boldsymbol{0}, \boldsymbol{I}_m \otimes \boldsymbol{I}_d)$ is a matrix of zero mean Gaussian distributed random variables. The covariance matrix $\boldsymbol{I}_m \otimes \boldsymbol{I}_d$ is an $m \times m$ diagonal block matrix of $d \times d$ identity matrices $\boldsymbol{I}_d$, that is,

$$\boldsymbol{I}_m \otimes \boldsymbol{I}_d = \begin{pmatrix} \boldsymbol{I}_d & \boldsymbol{0} & \ldots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_d & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{I}_d \end{pmatrix}. \tag{2.3.15}$$

Moreover, similarly to the vector case, we are able to define a transformed matrix Wiener process $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_t, t \in [0,\infty)\}$ using the above matrix Wiener process $\boldsymbol{W}$ as follows

$$\tilde{\boldsymbol{W}}_t = \boldsymbol{M}t + \boldsymbol{\Sigma}_1 \boldsymbol{W}_t \boldsymbol{\Sigma}_2^\top, \tag{2.3.16}$$

where $\boldsymbol{M}$ is a $d \times m$ matrix and $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are nonsingular $d \times d$ and $m \times m$ matrices, respectively. Values of such a matrix stochastic process can be obtained at times $t_i = i\Delta$ by the following recursive computation

$$\begin{aligned} \boldsymbol{W}_0 &= \boldsymbol{0} \\ \tilde{\boldsymbol{W}}_{t_{i+1}} &= \tilde{\boldsymbol{W}}_{t_i} + \boldsymbol{M}\Delta + \sqrt{\Delta}\tilde{\boldsymbol{N}}_{i+1}, \end{aligned} \tag{2.3.17}$$

for $i \in \{0, 1, \ldots\}$ and $\tilde{\boldsymbol{N}}_{i+1} \sim N_{d\times m}(\boldsymbol{0}, \boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)$. Here, the covariance matrix $\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1$ is again an $m \times m$ block matrix of the form

$$\boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1 = \begin{pmatrix} \sigma_{1,1}^2 \boldsymbol{\Sigma}_1 & \sigma_{1,2}^2 \boldsymbol{\Sigma}_1 & \ldots & \sigma_{1,m}^2 \boldsymbol{\Sigma}_1 \\ \sigma_{2,1}^2 \boldsymbol{\Sigma}_1 & \sigma_{2,2}^2 \boldsymbol{\Sigma}_1 & \ldots & \sigma_{2,m}^2 \boldsymbol{\Sigma}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{m,1}^2 \boldsymbol{\Sigma}_1 & \sigma_{m,2}^2 \boldsymbol{\Sigma}_1 & \ldots & \sigma_{m,m}^2 \boldsymbol{\Sigma}_1 \end{pmatrix}, \tag{2.3.18}$$

where $\boldsymbol{\Sigma}_1 = [\sigma_{i,j}^1]_{i,j}^d$ and $\boldsymbol{\Sigma}_2 = [\sigma_{i,j}^2]_{i,j}^m$.

Let us illustrate this matrix valued stochastic process for a 2×2 matrix case. In Fig. 2.3.2 we display a transformed matrix Wiener process $\tilde{\boldsymbol{W}}$, which was obtained from a standard 2×2 matrix Wiener process $\boldsymbol{W}$ by the following transformation

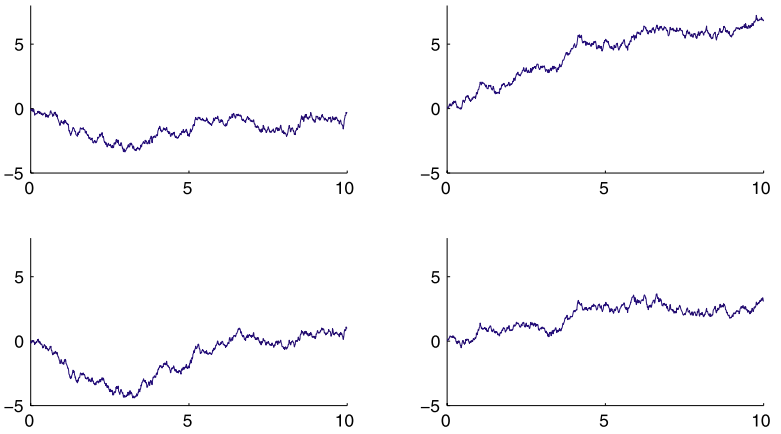

Fig. 2.3.2. 2×2 matrix Wiener process with independent elements in rows and dependent elements in columns

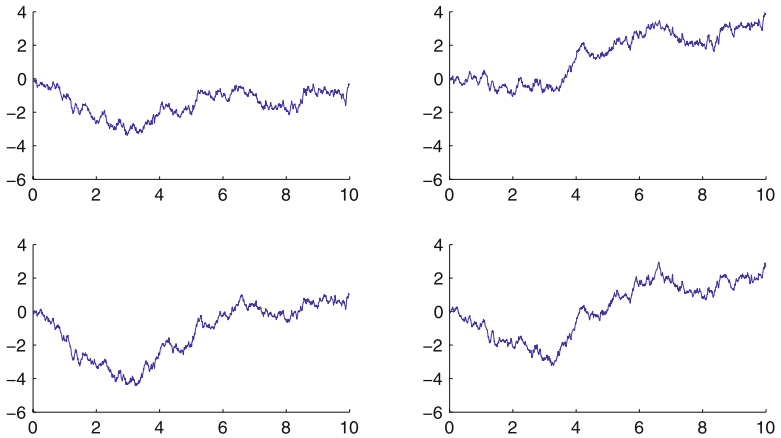

Fig. 2.3.3. 2×2 matrix Wiener process with both correlated rows and columns

$$\tilde{\boldsymbol{W}}_t = \boldsymbol{\Sigma}_1 \boldsymbol{W}_t \boldsymbol{I}^\top, \tag{2.3.19}$$

where $\boldsymbol{\Sigma}_1 = \boldsymbol{B}$ is as in (2.3.13) with correlation $\varrho = 0.8$. Note that in this case we obtain a matrix stochastic process whose rows have independent elements while its columns are formed by correlated Wiener processes. Similarly, in Fig. 2.3.3 we illustrate a 2×2 matrix transformed Wiener process $\tilde{\boldsymbol{W}}$, which was obtained from the standard 2×2 matrix Wiener process $\boldsymbol{W}$ by the transformation of the following form

$$\tilde{\boldsymbol{W}}_t = \boldsymbol{\Sigma}_1 \boldsymbol{W}_t \boldsymbol{\Sigma}_2^\top, \tag{2.3.20}$$

where both $\boldsymbol{\Sigma}_1 = \boldsymbol{B}$ and $\boldsymbol{\Sigma}_2 = \boldsymbol{B}$ are as in (2.3.13). For $\varrho = 0.8$ we note in Fig. 2.3.3 the correlation effect on the trajectories on both the elements of the columns and rows of such a 2×2 matrix transformed Wiener process.

Time Changed Wiener Processes

Instead of multiplying the time by some constant to scale the fluctuations of the Wiener paths one can introduce a flexible time dependent scaling by a, so called, time change. Let us now consider a vector of time changed standard independent Wiener processes $\boldsymbol{W}_{\varphi(t)} = (W^1_{\varphi(t)}, \ldots, W^d_{\varphi(t)})^\top$. Given the time discretization $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$, with time step size $\Delta > 0$ we obtain this time changed standard Wiener process by the following iterative formula

$$\begin{aligned} \boldsymbol{W}_{\varphi(0)} &= \boldsymbol{0} \\ \boldsymbol{W}_{\varphi(t_{i+1})} &= \boldsymbol{W}_{\varphi(t_i)} + \sqrt{\varphi(t_{i+1}) - \varphi(t_i)}\boldsymbol{N}_{i+1}, \end{aligned} \tag{2.3.21}$$

where the vector $\boldsymbol{N}_{i+1} \sim N_d(\boldsymbol{0}, \boldsymbol{I})$ is formed by independent standard Gaussian random variables. Obviously, it is possible to apply different time changes to different elements of the vector $\boldsymbol{W}$. For instance to prepare the representation of Ornstein-Uhlenbeck processes, let us define

$$\varphi_j(t) = \frac{b_j^2}{2c_j}(e^{2c_j t} - 1) \tag{2.3.22}$$

for $t \in [0, \infty)$, $b_j > 0, c_j > 0$ and $j \in \{1, 2, \ldots, d\}$, see (2.2.30). Then the elements of the vector $\boldsymbol{W}_{\varphi(t)}$ are such that

$$W^j_{\varphi_j(t_{i+1})} - W^j_{\varphi_j(t_i)} \sim N(0, \varphi_j(t_{i+1}) - \varphi_j(t_i)), \tag{2.3.23}$$

where $W^j_{\varphi_j(0)} = 0$, $j \in \{1, 2, \ldots, d\}$ and $i \in \{0, 1, \ldots\}$.

In order to obtain a time changed vector Wiener process, whose elements are correlated time changed Wiener processes, it is sufficient to define a new vector $\tilde{\boldsymbol{W}} = \{\tilde{\boldsymbol{W}}_{\varphi(t)} = (\tilde{W}^1_{\varphi(t)}, \ldots, \tilde{W}^d_{\varphi(t)})^\top, t \in [0, \infty)\}$ by the following transformation

$$\tilde{\boldsymbol{W}}_{\varphi(t)} = \boldsymbol{B}\boldsymbol{W}_{\varphi(t)}, \tag{2.3.24}$$

where $\boldsymbol{B}$ is a $d \times m$-matrix of coefficients and $\boldsymbol{W}_{\varphi(t)} = (W^1_{\varphi(t)}, \ldots, W^m_{\varphi(t)})^\top$ is an m-dimensional time changed Wiener process with independent components as in (2.3.23).

Additionally, let us define a $d \times m$ standard time changed matrix Wiener process $\boldsymbol{W} = \{\boldsymbol{W}_{\varphi(t)} = [W^{j,k}_{\varphi(t)}]^{d,m}_{j,k=1}, t \in [0, \infty)\}$. Here, the independent elements of the matrix $\boldsymbol{W}_{\varphi(t)}$ are such that

$$W^{j,k}_{\varphi_{j,k}(t_{i+1})} - W^{j,k}_{\varphi_{j,k}(t_i)} \sim N\left(0, \varphi_{j,k}(t_{i+1}) - \varphi_{j,k}(t_i)\right), \tag{2.3.25}$$

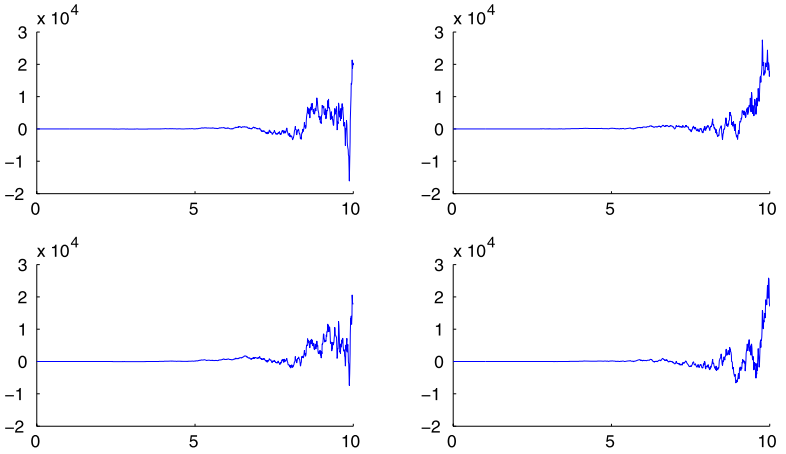

Fig. 2.3.4. Matrix valued time changed Wiener process

where $W^{k,j}_{\varphi_{k,j}(0)} = 0$, $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$ and $j \in \{1, 2, \ldots, d\}$, $k \in \{1, 2, \ldots, m\}$. Again, we may, for instance, define the (j,k)-th time transformation by

$$\varphi_{j,k}(t) = \frac{b_{j,k}^2}{2c_{j,k}} \left(e^{2c_{j,k}t} - 1\right) \tag{2.3.26}$$

for $t \in [0, \infty)$, $b_{j,k} > 0, c_{j,k} > 0$, and $j \in \{1, 2, \ldots, d\}$, $k \in \{1, 2, \ldots, m\}$. In order to obtain the time changed matrix Wiener process with correlated elements we can use the formula (2.3.16).

In Fig. 2.3.4 we display a matrix time changed Wiener process for $d = m = 2$ with the covariance matrix $\boldsymbol{I} \otimes \boldsymbol{\Sigma}_1$, where $\boldsymbol{\Sigma}_1$ is as in (2.3.13), $\varrho = 0.8$ and the parameters in the time change equal $b_{j,k} = \sqrt{2}$ and $c_{j,k} = 1$ for $j, k = \{1, 2\}$. That is, the same time change is applied to each of the elements of this matrix Wiener process. Namely, we construct $\tilde{\boldsymbol{W}}$ by the relation

$$\tilde{\boldsymbol{W}}_{\varphi(t)} = \boldsymbol{\Sigma}_1 \boldsymbol{W}_{\varphi(t)} \boldsymbol{I}^\top. \tag{2.3.27}$$

In this case we obtain a matrix time changed Wiener process $\tilde{\boldsymbol{W}}$ whose rows have independent elements, while columns have dependent elements.

Multi-dimensional OU-Processes

Let us now consider multi-dimensional Ornstein-Uhlenbeck (OU)-processes, covering both vector and matrix valued OU-processes, see Sect. 1.7. We will here construct the multi-dimensional OU-process as a time changed and scaled multi-dimensional Wiener process. Note that given the following two functions

$$s_t = \exp\{-ct\} \quad \text{and} \quad \varphi(t) = \frac{b^2}{2c}(e^{2ct} - 1) \tag{2.3.28}$$

for $t \in [0, \infty)$, $b, c > 0$, the scalar OU-process $Y = \{Y_t, t \in \Re\}$ can be represented in terms of a time changed and scaled scalar Wiener process, that is

$$Y_t = s_t W_{\varphi(t)}, \tag{2.3.29}$$

where $W = \{W_\varphi, \varphi \geq 0\}$ is a standard Wiener process in φ-time. By Itô's formula we obtain

$$\begin{aligned} dY_t &= W_{\varphi(t)}\, ds_t + s_t\, dW_{\varphi(t)} = -\frac{Y_t}{s_t}\, cs_t\, dt + s_t\, \frac{b}{s_t}\, d\tilde{W}_t \\ &= -cY_t\, dt + b\, d\tilde{W}_t, \end{aligned} \tag{2.3.30}$$

where $dW_{\varphi(t)} = \frac{b}{s_t}\, d\tilde{W}_t$, with $\tilde{W}$ denoting a standard Wiener process in t-time. Thus, by a straightforward time change and an application of the Itô formula one obtains a mean-reverting OU-process out of a basic Wiener process.

It is also straightforward to obtain a vector OU-process by

$$\boldsymbol{Y}_t = s_t \boldsymbol{W}_{\varphi(t)}, \tag{2.3.31}$$

that is,

$$Y_t^j = s_t^j W^j_{\varphi_j(t)} \tag{2.3.32}$$

for $j \in \{1, 2, \ldots, d\}$ and $t \geq 0$. The generalization to a matrix OU-process is obvious. The construction of this process starts by forming a $d \times m$ matrix time changed Wiener process and then scaling each element of this matrix by a function $s_t^{j,k}$ for $j \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$. Hence, the elements of such a matrix can be expressed by

$$Y_t^{j,k} = s_t^{j,k} W^{j,k}_{\varphi_{j,k}(t)} \tag{2.3.33}$$

for $j \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$.

We illustrate in Fig. 2.3.5 the matrix OU-process obtained from the matrix time changed Wiener process in Fig. 2.3.4 by use of formula (2.3.33). Since, the matrix time changed Wiener process has correlated rows and independent columns, the OU-process in Fig. 2.3.5 shares this feature.

Multi-dimensional SR-Processes via OU-Processes

Now let us consider δ OU-processes, that is

$$dX_t^i = -cX_t^i dt + b\, dW_t^i \tag{2.3.34}$$

for $t \in [0, \infty)$, with $X_0^i = x_0^i$, $c, b \in \Re$ and independent standard Wiener processes W^i for $i \in \{1, 2, \ldots, \delta\}$, $\delta \in \{1, 2, \ldots\}$. The square of such an OU-process has the Itô differential

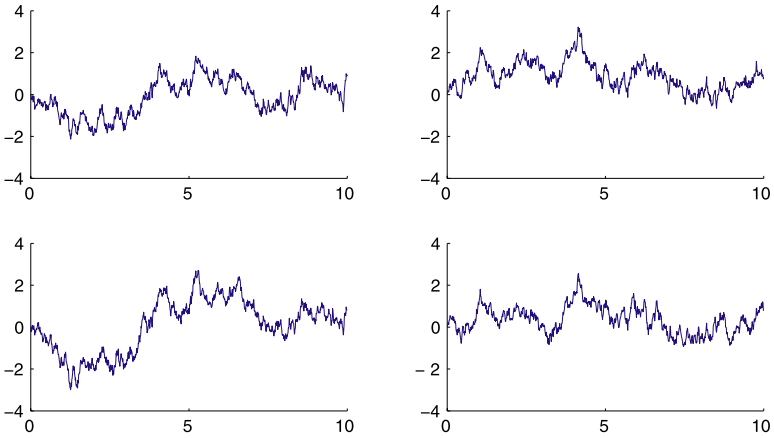

Fig. 2.3.5. Matrix valued Ornstein-Uhlenbeck process

$$d(X_t^i)^2 = (b^2 - 2c(X_t^i)^2)\,dt + 2bX_t^i\,dW_t^i, \tag{2.3.35}$$

for $t \in [0, \infty)$ and $i \in \{1, 2, \ldots, \delta\}$. Furthermore, we can form the sum of the δ squared OU-processes, that is,

$$Y_t = \sum_{i=1}^{\delta} (X_t^i)^2 \tag{2.3.36}$$

for $t \in [0, \infty)$. The SDE for Y_t is derived to be

$$dY_t = \sum_{i=1}^{\delta} \left(b^2 - 2c(X_t^i)^2\right)dt + 2b\sum_{i=1}^{\delta} X_t^i\,dW_t^i \tag{2.3.37}$$

for $t \in [0, \infty)$. In order to simplify the above SDE we introduce another Wiener process $\bar{W} = \{\bar{W}_t, t \in [0, \infty)\}$ defined as

$$\bar{W}_t = \int_0^t d\bar{W}_s = \sum_{i=1}^{\delta} \int_0^t \frac{X_s^i}{\sqrt{Y_s}}\,dW_s^i \tag{2.3.38}$$

for $t \in [0, \infty)$. It can be shown that the quadratic variation of $\bar{W}$ equals

$$[\bar{W}]_t = \int_0^t \sum_{i=1}^{\delta} \frac{(X_s^i)^2}{Y_s}\,ds = t. \tag{2.3.39}$$

Hence, by the Lévy theorem, see Theorem 1.3.3, we see that $\bar{W}$ is a standard Wiener process. Therefore, we obtain an equivalent SDE for the square root process Y in the form

$$dY_t = (\delta b^2 - 2cY_t)\,dt + 2b\sqrt{Y_t}\,d\bar{W}_t \tag{2.3.40}$$

for $t \in [0, \infty)$ with $Y_0 = \sum_{i=1}^{\delta} (x_0^i)^2$. Note that this process is a SR-process of dimension $\delta \in \{1, 2, \ldots\}$. It is well-known that for $\delta = 1$ the value Y_t can reach zero and is reflected at this boundary. For $\delta \in \{2, 3, \ldots\}$ the process never reaches zero for $Y_0 > 0$, see Revuz & Yor (1999).

Matrix Valued Squares of OU-Processes

Kendall (1989) and Bru (1991) studied the matrix generalization for squares of OU-processes. Denote by $\boldsymbol{X}_t$ a $\delta \times m$ matrix solution of the SDE

$$d\boldsymbol{X}_t = -c\boldsymbol{X}_t\,dt + b\,d\boldsymbol{W}_t, \tag{2.3.41}$$

for $t \geq 0$, with $\boldsymbol{X}_0 = \boldsymbol{x}_0$. Here $\boldsymbol{W}_t$ is a $\delta \times m$ matrix Wiener process and $\boldsymbol{x}_0$ is a $\delta \times m$ deterministic initial matrix; $b, c \in \Re$. By setting

$$\boldsymbol{S}_t = \boldsymbol{X}_t^\top \boldsymbol{X}_t, \quad \boldsymbol{s}_0 = \boldsymbol{x}_0^\top \boldsymbol{x}_0 \tag{2.3.42}$$

and denoting $d\tilde{\boldsymbol{W}}_t = \sqrt{\boldsymbol{S}_t^{-1}}\boldsymbol{X}_t^\top d\boldsymbol{W}_t$ we obtain an $m \times m$ matrix Wiener process $\tilde{\boldsymbol{W}}_t$. Note that the elements of $\tilde{\boldsymbol{W}}_t$ can be correlated. Then $\boldsymbol{S}_t$ solves the SDE

$$d\boldsymbol{S}_t = (\delta b^2 \boldsymbol{I} - 2c\boldsymbol{S}_t)\,dt + b(\sqrt{\boldsymbol{S}_t}\,d\tilde{\boldsymbol{W}}_t + d\tilde{\boldsymbol{W}}_t^\top \sqrt{\boldsymbol{S}_t}) \tag{2.3.43}$$

for $t \geq 0$, $\boldsymbol{S}_0 = \boldsymbol{s}_0$, where $\boldsymbol{I}$ is the identity matrix. Here $\boldsymbol{S}_t$ corresponds to a continuous-time process of stochastic, symmetric, positive definite matrices, while $\sqrt{\boldsymbol{S}_t}$ is the positive symmetric square root of the matrix $\boldsymbol{S}_t$, see Gouriéroux & Sufana (2004). Furthermore, $\boldsymbol{S}_t^{-1}$ is the inverse of the symmetric positive definite $m \times m$ matrix $\boldsymbol{S}_t$ and $\sqrt{\boldsymbol{S}_t^{-1}}$ its square root.

The matrix SR-process $\boldsymbol{S}$ can be simulated given the above matrix OU-process and using the transform (2.3.42). Note that for $m = 1$ the transform (2.3.42) simplifies to (2.3.36). We illustrate in Fig. 2.3.6 the matrix SR-process obtained from the matrix OU-process from Fig. 2.3.5. Note that not all elements of such a matrix always remain positive. The elements $S^{1,2}$ and $S^{2,1}$ are identical and, in general, need not be positive. Most importantly, the diagonal elements $S^{1,1}$ and $S^{2,2}$ are correlated SR-processes, which are always positive.

Multi-dimensional Squared Bessel Processes

Another important stochastic process in financial and other modeling is the squared Bessel process (BESQ$_x^\delta$) $X = \{X_\varphi,\ \varphi \in [\varphi_0, \infty)\}$, $\varphi_0 \geq 0$, of dimension $\delta \geq 0$, see Revuz & Yor (1999). We present this scalar process here, since the solution of the corresponding SDE can be simulated exactly in a

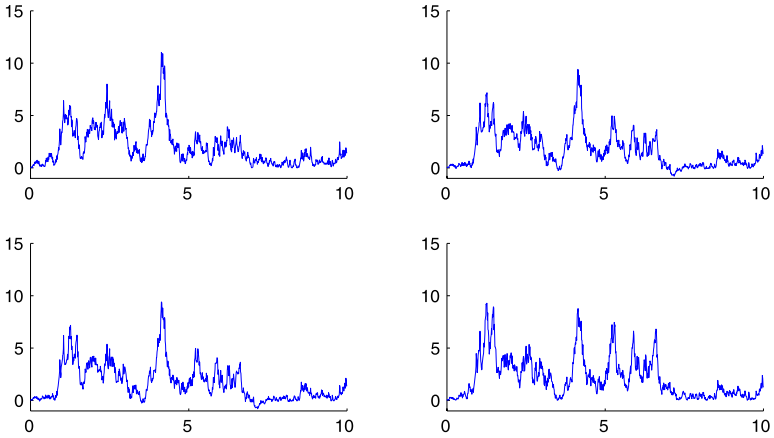

Fig. 2.3.6. Matrix valued square root process

convenient way for the case when the dimension of this process is an integer $\delta \in \{1, 2, \ldots\}$. This process can be described by the following SDE

$$dX_\varphi = \delta \, d\varphi + 2 \sqrt{|X_\varphi|} \, dW_\varphi \tag{2.3.44}$$

for $\varphi \in [\varphi_0, \infty)$ with $X_{\varphi_0} = x \geq 0$, where $W = \{W_\varphi, \varphi \in [\varphi_0, \infty)\}$ is a standard Wiener process in φ-time starting at the initial φ-time, $\varphi = \varphi_0$, at zero. This means, for $\varphi \in [\varphi_0, \infty)$ one has the increment in the quadratic variation of W as

$$[W]_\varphi - [W]_{\varphi_0} = \varphi - \varphi_0$$

for all $\varphi \in [\varphi_0, \infty)$. Furthermore, if we fix the behavior of X_φ at zero as reflection, then the absolute sign under the square root in (2.3.44) can be removed, and X_φ remains nonnegative and has a unique strong solution, see Revuz & Yor (1999).

It is a useful fact that for $\delta \in \{1, 2, \ldots\}$ and $x \geq 0$ the dynamics of a BESQ_x^δ process X can be expressed as the sum of the squares of δ independent Wiener processes $W^1, W^2, \ldots, W^\delta$ in φ-time, which start at time $\varphi = 0$ in $w^1 \in \Re, w^2 \in \Re, \ldots w^\delta \in \Re$, respectively, such that

$$x = \sum_{k=1}^{\delta} (w^k)^2. \tag{2.3.45}$$

We can now construct a solution of (2.3.44) as follows

$$X_\varphi = \sum_{k=1}^{\delta} (w^k + W_\varphi^k)^2 \tag{2.3.46}$$

for $\varphi \in [0, \infty)$. Applying the Itô formula we obtain

$$dX_\varphi = \delta d\varphi + 2 \sum_{k=1}^{\delta} (w^k + W_\varphi^k) dW_\varphi^k \tag{2.3.47}$$

for $\varphi \in [0, \infty)$ with

$$X_0 = \sum_{k=1}^{\delta} (w^k)^2 = x. \tag{2.3.48}$$

Furthermore, by setting

$$dW_\varphi = |X_\varphi|^{-\frac{1}{2}} \sum_{k=1}^{\delta} (w^k + W_\varphi^k) dW_\varphi^k \tag{2.3.49}$$

we satisfy with (2.3.46) the SDE (2.3.44). Note that we have for W_φ the quadratic variation

$$[W]_\varphi = \int_0^\varphi \frac{1}{X_s} \sum_{k=1}^{\delta} (w^k + W_s^k)^2 ds = \varphi. \tag{2.3.50}$$

Hence, by the Lévy theorem, see Theorem 1.3.3, W_φ is a Wiener process in φ-time.

Wishart Process

The matrix generalization of a squared Bessel process is a Wishart process, see Bru (1991). The $m \times m$ matrix valued Wishart process with dimension $\delta \in \{1, 2, \ldots\}$ is the matrix process $\boldsymbol{S} = \{\boldsymbol{S}_t, t \geq 0\}$ with

$$\boldsymbol{S}_t = \boldsymbol{W}_t^\top \boldsymbol{W}_t \tag{2.3.51}$$

and initial matrix $\boldsymbol{s}_0 = \boldsymbol{W}_0^\top \boldsymbol{W}_0$ for $t \in \Re^+$, where $\boldsymbol{W}_t$ is the value at time $t \geq 0$ of a $\delta \times m$ matrix Wiener process. Itô calculus applied to the relation (2.3.51) results in the following SDE

$$d\boldsymbol{S}_t = \delta \boldsymbol{I} dt + d\boldsymbol{W}_t^\top \boldsymbol{W}_t + \boldsymbol{W}_t^\top d\boldsymbol{W}_t, \tag{2.3.52}$$

where $\boldsymbol{I}$ is the $m \times m$ identity matrix. It can be shown that $\tilde{\boldsymbol{W}}_t$ expressed by

$$d\tilde{\boldsymbol{W}}_t = \left(\sqrt{\boldsymbol{S}_t}\right)^{-1} \boldsymbol{W}_t^\top d\boldsymbol{W}_t \tag{2.3.53}$$

is an $m \times m$ matrix Wiener process. Here $\sqrt{\boldsymbol{S}_t}$ represents the symmetric positive square root of $\boldsymbol{S}_t$, while $\left(\sqrt{\boldsymbol{S}_t}\right)^{-1}$ is the inverse of the matrix $\sqrt{\boldsymbol{S}_t}$. Note also that

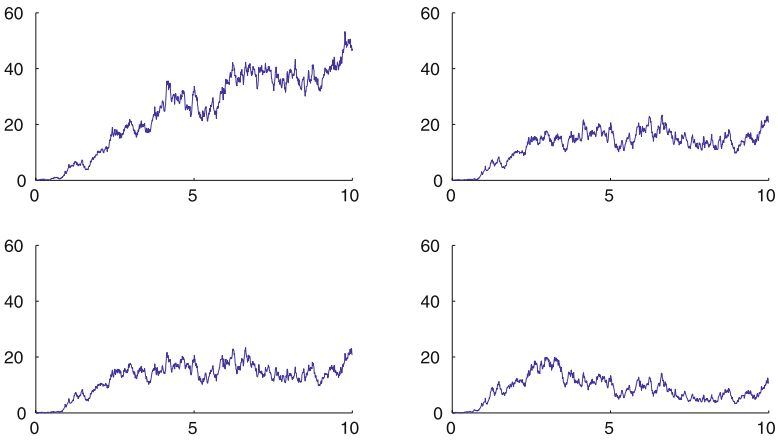

Fig. 2.3.7. Wishart process

$$d\tilde{\boldsymbol{W}}_t^\top = d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\left(\sqrt{\boldsymbol{S}_t} \right)^{-1} \right)^\top = d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\left(\sqrt{\boldsymbol{S}_t} \right)^\top \right)^{-1} \tag{2.3.54}$$
$$= d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\sqrt{\boldsymbol{S}_t^\top} \right)^{-1} = d\boldsymbol{W}_t^\top \boldsymbol{W}_t \left(\sqrt{\boldsymbol{S}_t} \right)^{-1},$$

since $\boldsymbol{S}_t$ is a symmetric matrix. Therefore, (2.3.52) can be rewritten in the following form

$$d\boldsymbol{S}_t = \delta \boldsymbol{I} dt + \sqrt{\boldsymbol{S}_t} d\tilde{\boldsymbol{W}}_t + d\tilde{\boldsymbol{W}}_t^\top \sqrt{\boldsymbol{S}_t} \tag{2.3.55}$$

for $t \in \Re^+$.

In Fig. 2.3.2 we showed the trajectories of the elements of a 2×2 matrix Wiener process. Now, in Fig. 2.3.7 we plot a 2×2 Wishart process of dimension $\delta = 2$ obtained from the Wiener process in Fig. 2.3.2. Recall that the matrix Wiener process in this example was obtained by assuming the covariance matrix $\boldsymbol{I} \otimes \boldsymbol{\Sigma}_1$, where $\boldsymbol{\Sigma}_1$ is as in (2.3.13).

SR-Processes via Squared Bessel Processes

Using squared Bessel processes one can derive SR-processes by certain transformations. For this reason let $c : [0, \infty) \to \Re$ and $b : [0, \infty) \to \Re$ be given deterministic functions of time. We introduce the exponential

$$s_t = s_0 \exp \left\{ \int_0^t c_u \, du \right\} \tag{2.3.56}$$

and the φ-time

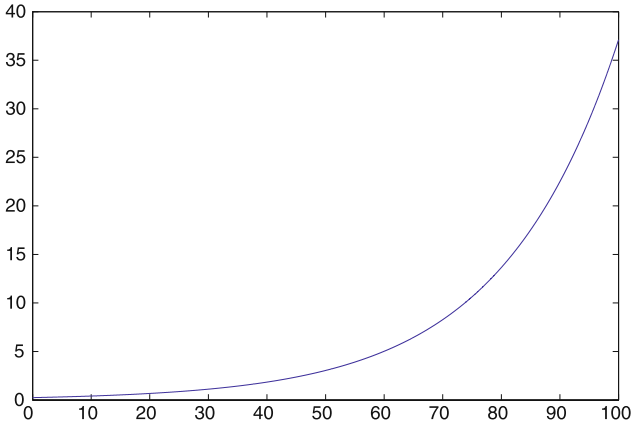

Fig. 2.3.8. Time $\varphi(t)$ against time t

$$\varphi(t) = \varphi(0) + \frac{1}{4}\int_0^t \frac{b_u^2}{s_u}\,du \tag{2.3.57}$$

for $t \in [0, \infty)$ and $s_0 > 0$ dependent on t-time. Note that we have an explicit representation for the function $\varphi(t)$ in the case of constant parameters $b_t = \bar{b} \neq 0$ and $c_t = \bar{c} \neq 0$, where

$$\varphi(t) = \varphi(0) + \frac{\bar{b}^2}{4\bar{c}s_0}(1 - \exp\{-\bar{c}t\}) \tag{2.3.58}$$

for $t \in [0, \infty)$ and $s_0 > 0$. Furthermore, if $\varphi(0) = -\frac{\bar{b}^2}{4\bar{c}s_0}$, this function simply equals

$$\varphi(t) = -\frac{\bar{b}^2}{4\bar{c}s_0}\exp\{-\bar{c}t\} \tag{2.3.59}$$

for $t \in [0, \infty)$, $s_0 > 0$, $\bar{b} \neq 0$ and $\bar{c} \neq 0$. We show the function $\varphi(t)$ in Fig. 2.3.8 for the choice of $\bar{b} = 1$, $\bar{c} = -0.05$, $s_0 = 20$ and $\varphi(0) = -\frac{\bar{b}^2}{4\bar{c}s_0} = 0.25$. The function $\varphi(t)$ is a time transformation, which when applied to the BESQ_x^δ yields the following expected value of the time transformed squared Bessel process

$$E(X_{\varphi(t)}|\mathcal{A}_{\varphi(0)}) = X_{\varphi(0)} + \delta(\varphi(t) - \varphi(0)), \tag{2.3.60}$$

for $t \in [0, \infty)$. Note also that for constant parameters $b_t = \bar{b} \neq 0$, $c_t = \bar{c} \neq 0$ and $X_{\varphi(0)} = -\frac{\delta\bar{b}^2}{4\bar{c}s_0}$ this expected value simplifies to

$$E(X_{\varphi(t)}|\mathcal{A}_{\varphi(0)}) = -\frac{\delta\bar{b}^2}{4\bar{c}s_0}\exp\{-\bar{c}t\} \tag{2.3.61}$$

for $t \in [0, \infty)$.

Given a squared Bessel process X of dimension $\delta > 0$ and using our previous notation we introduce the SR-process $Y = \{Y_t, t \geq 0\}$ of dimension $\delta > 0$ with

$$Y_t = s_t \, X_{\varphi(t)} \tag{2.3.62}$$

in dependence on time $t \geq 0$, see also Delbaen & Shirakawa (2002).

Further, by (2.3.44), (2.3.62), (2.3.56) and (2.3.57) and the Itô formula we can express (2.3.62) in terms of the SDE

$$dY_t = \Big(\frac{\delta}{4} \, b_t^2 + c_t \, Y_t\Big) dt + b_t \, \sqrt{Y_t} \, dU_t \tag{2.3.63}$$

for $t \in [0, \infty)$, $Y_0 = s_0 X_{\varphi(0)}$ and

$$dU_t = \sqrt{\frac{4 s_t}{b_t^2}} dW_{\varphi(t)}.$$

Note that U_t forms by the Lévy theorem, see Theorem 1.3.3, a Wiener process, since

$$[U]_t = \int_0^t \frac{4 s_z}{b_z^2} d\varphi(z) = t. \tag{2.3.64}$$

The same time-change formula applies in the more general matrix case. Given a Wishart process $\boldsymbol{X}$ it can be shown, that the matrix square root process can be obtained from the Wishart process by the following transformation

$$\boldsymbol{Y}_t = s_t \boldsymbol{X}_{\varphi(t)}, \tag{2.3.65}$$

where s_t and $\varphi(t)$ are as in (2.3.56) and (2.3.57), respectively. By (2.3.55), (2.3.65), (2.3.56) and (2.3.57) and the Itô formula we can express (2.3.65) in terms of the matrix SDE

$$d\boldsymbol{Y}_t = \left(\frac{\delta}{4} b_t^2 \boldsymbol{I} + c_t \boldsymbol{Y}_t\right) dt + \frac{b_t}{2} \left(\sqrt{\boldsymbol{Y}_t} d\boldsymbol{U}_t + d\boldsymbol{U}_t^\top \sqrt{\boldsymbol{Y}_t}\right) \tag{2.3.66}$$

for $t \in [0, \infty)$, $\boldsymbol{Y}_0 = s_0 \boldsymbol{X}_{\varphi(0)}$ and where $d\boldsymbol{U}_t = \sqrt{\frac{4 s_t}{b_t^2}} d\boldsymbol{W}_{\varphi(t)}$ is the differential of a matrix Wiener process.

In Fig. 2.3.9 we display the trajectory of the elements of a 2×2 matrix time changed Wishart process $\boldsymbol{X}_{\varphi(t)}$ in log-scale. Here the off-diagonal elements do not show any value for the time periods when the argument of the logarithm becomes negative. Indeed we see in Fig. 2.3.9 that the off-diagonal elements have such negative values near the time $t = 7$. This is not the case for the diagonal elements which are of main interest in most applications. In Fig. 2.3.10 we show the corresponding trajectory of a 2×2 matrix SR-process obtained from the time changed Wishart process by the use of formula (2.3.65). Note that this matrix SR-process is identical to the matrix SR-process in Fig. 2.3.6 obtained via squares of OU-processes.

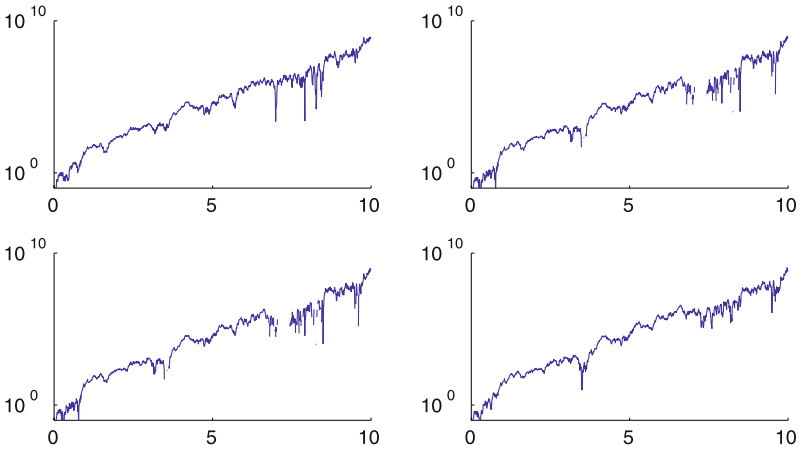

Fig. 2.3.9. Time changed Wishart process in log-scale

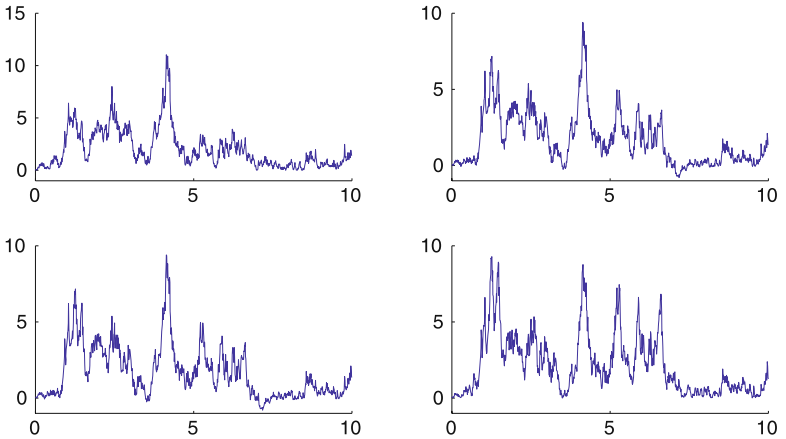

Fig. 2.3.10. Matrix valued square root process

Multi-dimensional Affine Processes

Let us now further transform the above obtained multi-dimensional SR-process in order to obtain multi-dimensional affine processes, see Sect. 1.6. These processes have affine, that is linear drift and linear squared diffusion coefficients. In order to obtain members of this class of multi-dimensional processes we can simply shift the multi-dimensional SR-process by a nonnegative, differentiable function of time $a : [0, \infty) \to [0, \infty)$ characterized by its derivative

$$a_t' = \frac{da_t}{dt} \tag{2.3.67}$$

for $t \in [0, \infty)$ with $a_0 \in [0, \infty)$. More precisely, we define the process $\boldsymbol{R} = \{\boldsymbol{R}_t, t \in [0, \infty)\}$ such that

$$\boldsymbol{R}_t = \boldsymbol{Y}_t + a_t \boldsymbol{I} \tag{2.3.68}$$

for $t \in [0, \infty)$. It is also possible to obtain more general affine processes by shifting the matrix valued SR-process by a matrix $\boldsymbol{A}_t$ of nonnegative differentiable functions of the type (2.3.67). That is

$$\boldsymbol{R}_t = \boldsymbol{Y}_t + \boldsymbol{A}_t \tag{2.3.69}$$

for $t \in [0, \infty)$. In this case $\boldsymbol{R}_t$ solves the following matrix SDE

$$d\boldsymbol{R}_t = \left(\frac{\delta}{4} b_t^2 \boldsymbol{I} + \boldsymbol{A}_t' - c_t \boldsymbol{A}_t + c_t \boldsymbol{R}_t\right) dt + \frac{b_t}{2} \left(\sqrt{\boldsymbol{R}_t - \boldsymbol{A}_t} d\tilde{\boldsymbol{W}}_t + d\tilde{\boldsymbol{W}}_t^\top \sqrt{\boldsymbol{R}_t - \boldsymbol{A}_t}\right), \tag{2.3.70}$$

for $t \in [0, \infty)$. Here $\boldsymbol{A}_t'$ denotes the matrix of the derivatives of the type (2.3.67) for the shifts of each element. In principle, we applied here the Itô formula to the equation (2.3.69).

Multi-dimensional Geometric Ornstein-Uhlenbeck Processes

The Itô formula provides a general tool to generate a world of exact solutions of SDEs based on functions of the solutions of those processes we have already considered. As an example, let us generate explicit solutions for a geometric OU-process. Here each element of a matrix valued OU-process is simply exponentiated. That is, denoting by $\boldsymbol{Y}_t = [Y_t^{j,k}]_{j,k=1}^{d,m}$ the corresponding $d \times m$ matrix geometric OU-process value at time t and by $\boldsymbol{X}_t = [X_t^{j,k}]_{j,k=1}^{d,m}$ the $d \times m$ matrix OU-process value. We obtain the elements of the matrix $\boldsymbol{Y}_t$ by

$$Y_t^{j,k} = \exp\{X_t^{j,k}\} \tag{2.3.71}$$

for $t \in [0, \infty)$ and $j \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$.

In Fig. 2.3.11 we illustrate a 2×2 matrix geometric OU-process obtained from the matrix OU-process in Fig. 2.3.5 by application of (2.3.71) to each of its elements. More complex applications of the Itô formula generating exact solutions will be considered in the next section.

Multi-dimensional SDEs Driven by Lévy Processes

So far in this section we have considered the exact simulation of solutions of multi-dimensional SDEs driven by vector or matrix Wiener processes. However, the simulation methods described here can be adapted to multi-dimensional SDEs which are driven by more general vector or matrix valued

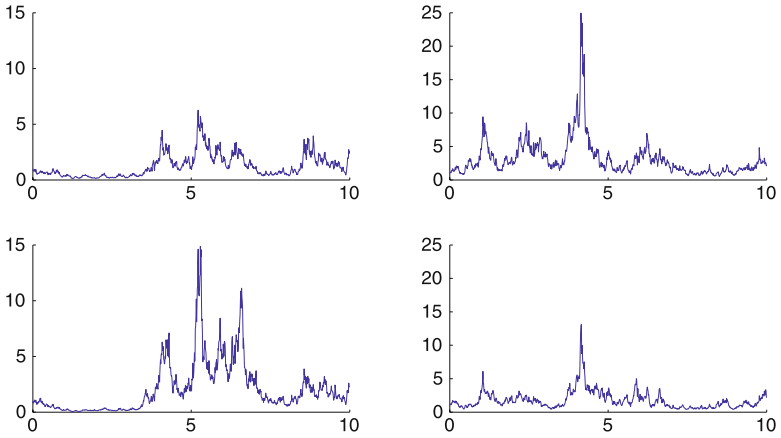

Fig. 2.3.11. Matrix valued geometric OU-process

Lévy processes, see Sect. 1.7. In principle, one can substitute the Wiener processes by some Lévy processes.

Since Lévy processes have independent stationary increments, it is possible to construct paths of a wide range of d-dimensional Lévy processes $\boldsymbol{L} = \{\boldsymbol{L}_t, t \geq 0\}$ at given discretization times $t_i = i\Delta$, $i \in \{0, 1, \ldots\}$, with fixed time step size $\Delta > 0$. The distribution of the Lévy increments $\boldsymbol{L}_{t_{i+1}} - \boldsymbol{L}_{t_i}$, however, must be infinitely divisible for the process $\boldsymbol{L}$ to be the transition distribution of a Lévy process, see Sect. 2.2. One example of a family of infinitely divisible distributions is the generalized hyperbolic (GH) distribution, see for instance McNeil et al. (2005). This family of distributions yields variance gamma (VG) and normal inverse Gaussian (NIG) processes as special cases.

Simulation of the d-dimensional VG and NIG processes results from their representation as subordinated vector Wiener processes with drift. That is,

$$\boldsymbol{L}_t = \boldsymbol{a}V_t + \boldsymbol{B}\boldsymbol{W}_{V_t}, \tag{2.3.72}$$

for $t \in [0, \infty)$. Here $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^\top$ is a d-dimensional vector, $\boldsymbol{B}$ is a $d \times m$-matrix and $\boldsymbol{W} = \{\boldsymbol{W}_V = (W_V^1, W_V^2, \ldots, W_V^m)^\top, V \in [0, \infty)\}$ is a standard m-dimensional vector Wiener process in V time. When V is the gamma process or the inverse Gaussian process, we obtain the d-dimensional VG process and the NIG process, respectively.

We also define $d \times m$ matrix VG and NIG processes by

$$\boldsymbol{L}_t = \boldsymbol{M}V_t + \boldsymbol{\Sigma}_1\boldsymbol{W}_{V_t}\boldsymbol{\Sigma}_2, \tag{2.3.73}$$

where $\boldsymbol{M}$ is a $d \times m$ matrix and $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are nonsingular $d \times d$ and $m \times m$ matrices, respectively. Here $\boldsymbol{W} = \{\boldsymbol{W}_V = [W_V^{i,j}]_{i,j=1}^{d,m}, V \in [0, \infty)\}$ is

a standard $d \times m$ matrix Wiener process and V a gamma or inverse Gaussian process, respectively.

Processes of type (2.3.72) and (2.3.73) possess a number of useful properties because they are conditionally Gaussian. In particular, if one knows how to simulate the increments of the subordinator V, the values of $\boldsymbol{L}$ in (2.3.73) can be obtained at the discrete times $t_i = i\Delta$ by the following recursive computation

$$\begin{aligned} \boldsymbol{L}_0 &= \boldsymbol{0} \\ \boldsymbol{L}_{t_{i+1}} &= \boldsymbol{L}_{t_i} + \boldsymbol{M}\Delta V_{i+1} + \sqrt{\Delta V_{i+1}}\tilde{\boldsymbol{N}}_{i+1}, \end{aligned} \tag{2.3.74}$$

for $i \in \{0, 1, \dots, \}$ and $\tilde{\boldsymbol{N}}_{i+1} \sim \mathcal{N}_{d\times m}(\boldsymbol{0}, \boldsymbol{\Sigma}_2 \otimes \boldsymbol{\Sigma}_1)$. Here the covariance matrix $\boldsymbol{\Sigma}_1 \otimes \boldsymbol{\Sigma}_2$ is as in (2.3.18).

The VG process is obtained by (2.3.74), where $\Delta V_{i+1} \sim \kappa \mathrm{Ga}(\frac{\Delta}{\kappa}, 1)$ are gamma random variables for $i \in \{1, 2, \dots\}$, while the NIG process is obtained by (2.3.74) when $\Delta V_{i+1} \sim \mathrm{IGaussian}(\frac{\Delta^2}{\kappa}, \Delta)$ are inverse Gaussian random variables for $i \in \{1, 2, \dots\}$. Here the parameter κ is the variance of the subordinator V. See also Cont & Tankov (2004) who describe exact simulation of scalar VG and NIG processes. They describe also convenient algorithms for generators of gamma and inverse Gaussian random variables.

Since we can simulate the paths of such driving Lévy processes exactly it is possible to simulate solutions for the type of the above introduced SDEs when driven by Lévy noise. For instance, let us consider a Wishart process of dimension δ driven by a VG-process. That is, we consider the multi-dimensional SDE of the form

$$d\boldsymbol{S}_t = \delta \boldsymbol{I} dt + \sqrt{\boldsymbol{S}_t} d\boldsymbol{L}_t + d\boldsymbol{L}_t^\top \sqrt{\boldsymbol{S}_t} \tag{2.3.75}$$

for $t \in \Re^+$. In order to simulate this Wishart process, which may be driven by the VG-process $\boldsymbol{L}$, we first need to simulate a $\delta \times m$ matrix VG-process. Afterwards we obtain the $m \times m$ VG-Wishart process of dimension $\delta \in \{1, 2, \dots\}$ by setting $\boldsymbol{S}_t = \boldsymbol{L}_t^\top \boldsymbol{L}_t$, for $t \in \Re^+$.

In Fig. 2.3.12 we show a trajectory of a gamma process, which is always nondecreasing. Here we have chosen $\kappa = 1$. Moreover, in Fig. 2.3.13 we display a 2×2 matrix VG-process, with parameters $\boldsymbol{M} = \boldsymbol{0}$ and the covariance matrix $\boldsymbol{I} \otimes \boldsymbol{\Sigma}_1$, where $\boldsymbol{\Sigma}_1 = B$ is as in (2.3.13) with $\varrho = 0.8$. The subordinator is here chosen to be the gamma process illustrated in Fig 2.3.12. Additionally, in Fig. 2.3.14 we display the corresponding trajectory of the resulting 2×2 Wishart process of dimension $\delta = 2$.

The subordination methodology can be widely applied to generate trajectories of other matrix Lévy processes, for instance, matrix Lévy OU-processes and matrix Lévy-affine processes.

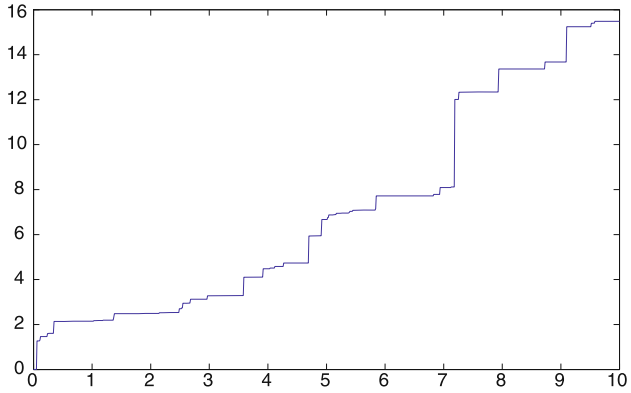

Fig. 2.3.12. Gamma process

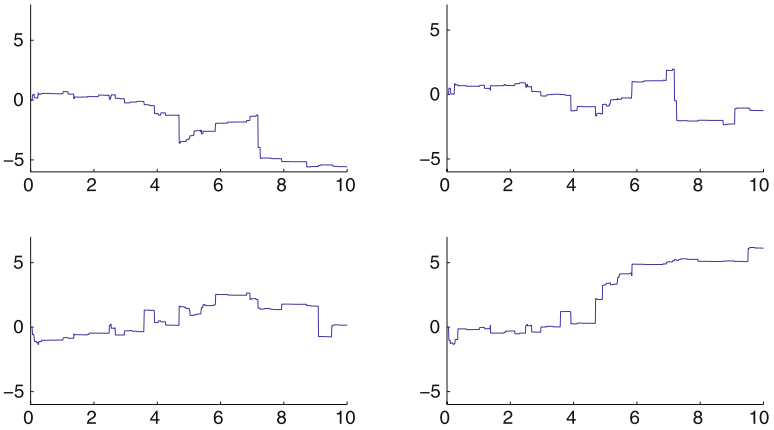

Fig. 2.3.13. Matrix VG-process

2.4 Functions of Exact Solutions

Another possibility to obtain a multi-dimensional SDE which is explicitly solvable, is by application of the Itô formula, see also Kloeden & Platen (1999). Let us illustrate this below for linear SDEs driven by a vector of standard Wiener processes.

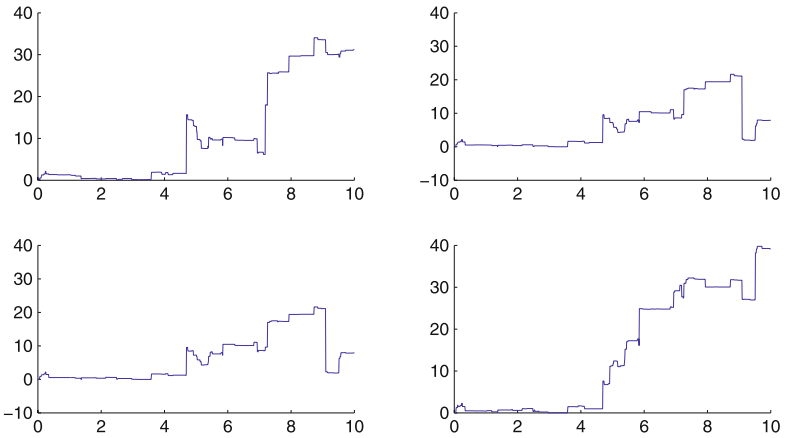

Fig. 2.3.14. Wishart process driven by the matrix VG-process in Fig. 2.3.13

Multi-dimensional Itô Formula Application

We consider an m-dimensional Wiener process $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, \ldots, W_t^m)^\top, t \in [0, \infty)\}$, a d-dimensional drift coefficient vector function $\boldsymbol{a} : [0, T] \times \Re^d \to \Re^d$ and a $d \times m$-matrix diffusion coefficient function $\boldsymbol{b} : [0, T] \times \Re^d \to \Re^{d \times m}$. In this framework we assume that we have already a general family of explicitly solvable d-dimensional SDEs given as

$$d\boldsymbol{X}_t = \boldsymbol{a}(t, \boldsymbol{X}_t)dt + \boldsymbol{b}(t, \boldsymbol{X}_t)d\boldsymbol{W}_t, \tag{2.4.1}$$

for $t \in [0, \infty)$, $\boldsymbol{X}_0 \in \Re^d$. This means that the kth component of (2.4.1) equals

$$dX_t^k = a^k(t, \boldsymbol{X}_t)dt + \sum_{j=1}^{m} b^{k,j}(t, \boldsymbol{X}_t)dW_t^j. \tag{2.4.2}$$

For a sufficiently smooth vector function $\boldsymbol{U} : [0, T] \times \Re^d \to \Re^k$ of the solution $\boldsymbol{X}_t$ of (2.4.1) we obtain a k-dimensional process

$$\boldsymbol{Y}_t = \boldsymbol{U}(t, \boldsymbol{X}_t). \tag{2.4.3}$$

The expression for its pth component, resulting from the application of the Itô formula, satisfies the SDE

$$\begin{aligned} dY_t^p = & \left(\frac{\partial U^p}{\partial t} + \sum_{i=1}^{d} a^i \frac{\partial U^p}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^{d} \sum_{l=1}^{m} b^{i,l} b^{j,l} \frac{\partial^2 U^p}{\partial x_i \partial x_j} \right) dt \\ & + \sum_{l=1}^{m} \sum_{i=1}^{d} b^{i,l} \frac{\partial U^p}{\partial x_i} dW_t^l, \end{aligned} \tag{2.4.4}$$

for $p \in \{1, 2, \ldots, k\}$, where the terms on the right-hand side of (2.4.4) are evaluated at $(t, \boldsymbol{X}_t)$. Obviously, also the paths of the solution of the SDE (2.4.4) can be exactly simulated since $\boldsymbol{X}_{t_i}$ can be obtained at all discretization points and by (2.4.3) the vector $\boldsymbol{Y}_{t_i}$ is just a function of the components of the vector $\boldsymbol{X}_{t_i}$.

Multi-dimensional Linear SDEs

In practice, the multi-dimensional Itô formula turns out to be a useful tool if one wants to construct solutions of certain multi-dimensional SDEs in terms of known solutions of other SDEs. Let us illustrate this with an example involving linear SDEs, where we rely on the results of Sect. 1.7. Recall from (1.7.20) that a d-dimensional linear SDE can be expressed in the form

$$d\boldsymbol{X}_t = (\boldsymbol{A}_t \, \boldsymbol{X}_t + \boldsymbol{\alpha}_t) \, dt + \sum_{l=1}^{m} \left(\boldsymbol{B}_t^l \, \boldsymbol{X}_t + \boldsymbol{\beta}_t^l \right) dW_t^l \tag{2.4.5}$$

for $t \geq 0$ with $\boldsymbol{X}_0 \in \Re^d$. Here $\boldsymbol{A}, \boldsymbol{B}^1, \boldsymbol{B}^2, \ldots, \boldsymbol{B}^m$ are deterministic $d \times d$ matrix functions of time and $\boldsymbol{\alpha}, \boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \ldots, \boldsymbol{\beta}^m$ are deterministic d-dimensional vector functions of time. It is possible to express according to (1.7.21) the solution of (2.4.5) in the form

$$\boldsymbol{X}_t = \boldsymbol{\Psi}_t \left(\boldsymbol{X}_0 + \int_0^t \boldsymbol{\Psi}_s^{-1} \left(\boldsymbol{\alpha}_s - \sum_{l=1}^{m} \boldsymbol{B}_s^l \, \boldsymbol{\beta}_s^l \right) ds + \sum_{l=1}^{m} \int_0^t \boldsymbol{\Psi}_s^{-1} \, \boldsymbol{\beta}_s^l \, dW_s^l \right). \tag{2.4.6}$$

Here $\boldsymbol{\Psi}_t$ is the $d \times d$ fundamental matrix satisfying $\boldsymbol{\Psi}_0 = \boldsymbol{I}$ and the homogeneous matrix SDE

$$d\boldsymbol{\Psi}_t = \boldsymbol{A}_t \, \boldsymbol{\Psi}_t \, dt + \sum_{l=1}^{m} \boldsymbol{B}_t^l \, \boldsymbol{\Psi}_t \, dW_t^l, \tag{2.4.7}$$

see (1.7.22). Unfortunately, it is not possible to solve (2.4.6) explicitly for its fundamental solution in its general form. However, if the matrices $\boldsymbol{A}, \boldsymbol{B}^1, \boldsymbol{B}^2, \ldots, \boldsymbol{B}^m$ are constant and commute, that is if

$$\boldsymbol{A}\boldsymbol{B}^l = \boldsymbol{B}^l \boldsymbol{A} \quad \text{and} \quad \boldsymbol{B}^l \boldsymbol{B}^k = \boldsymbol{B}^k \boldsymbol{B}^l \tag{2.4.8}$$

for all $k, l \in \{1, 2, \ldots, m\}$, then the explicit expression for the fundamental matrix solution is given as

$$\boldsymbol{\Psi}_t = \exp\left\{ \left(\boldsymbol{A} - \frac{1}{2} \sum_{l=1}^{m} (\boldsymbol{B}^l)^2 \right) t + \sum_{l=1}^{m} \boldsymbol{B}^l \, W_t^l \right\}, \tag{2.4.9}$$

where the exponential is interpreted elementwise, see (1.7.26). This allows to cover interesting multi-dimensional models that are relevant to finance and other areas of application.

Multi-dimensional Black-Scholes Model

We will now describe the multi-dimensional Black-Scholes model, which is the standard asset price model in finance. This model emerges from (2.4.5) by assuming that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}^l$ equal zero for all $l \in \{1, 2, \ldots, m\}$, see also (1.7.27)–(1.7.31). Denote by $\boldsymbol{S}_t$ a diagonal matrix with jth diagonal element S_t^j, $j \in \{1, 2, \ldots, d\}$, representing the jth asset price at time $t \in [0, \infty)$. Then the SDE for the jth Black-Scholes asset price S_t^j is defined by

$$dS_t^j = S_t^j \left(a_t^j dt + \sum_{k=1}^{d} b_t^{j,k} dW_t^k \right) \tag{2.4.10}$$

for $t \in [0, \infty)$ and $j \in \{1, 2, \ldots, d\}$, where a^j and $b^{j,k}$ are deterministic functions of time, see (1.7.13) and (1.7.31). Here $W^k, k \in \{1, 2, \ldots, d\}$, denote independent standard Wiener processes. To represent this SDE in matrix form we introduce the diagonal matrix $\boldsymbol{A}_t = [A_t^{i,j}]_{i,j=1}^d$ with

$$A_t^{i,j} = \begin{cases} a_t^j & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.4.11}$$

and diagonal matrix $\boldsymbol{B}_t^k = [B_t^{k,i,j}]_{i,j=1}^d$ with

$$B_t^{k,i,j} = \begin{cases} b_t^{j,k} & \text{for } i = j \\ 0 & \text{otherwise} \end{cases} \tag{2.4.12}$$

for $k, i, j \in \{1, 2, \ldots, d\}$ and $t \in [0, \infty)$. All these diagonal matrices commute in the sense of (2.4.8), therefore, we can write the SDE (2.4.10) as the matrix SDE

$$d\boldsymbol{S}_t = \boldsymbol{A}_t \boldsymbol{S}_t \, dt + \sum_{k=1}^{d} \boldsymbol{B}_t^k \boldsymbol{S}_t \, dW_t^k \tag{2.4.13}$$

for $t \in [0, \infty)$. Consequently, we obtain for the jth asset price the explicit solution

$$S_t^j = S_0^j \exp\left\{ \int_0^t \left(a_s^j - \frac{1}{2} \sum_{k=1}^{d} (b_t^{j,k})^2 \right) ds + \sum_{k=1}^{d} \int_0^t b_s^{j,k} dW_s^k \right\} \tag{2.4.14}$$

for $t \in [0, \infty)$ and $j \in \{1, 2, \ldots, d\}$. When taking the following exponential elementwise, the explicit solution of (2.4.10) can be expressed as

$$\boldsymbol{S}_t = \boldsymbol{S}_0 \exp\left\{ \int_0^t \left(\boldsymbol{A}_s - \frac{1}{2} \sum_{k=1}^{d} \left(\boldsymbol{B}_s^k \right)^2 \right) ds + \sum_{k=1}^{d} \int_0^t \boldsymbol{B}_s^k dW_s^k \right\} \tag{2.4.15}$$

for $t \geq 0$, see (1.7.31). If the appreciation rates and volatilities are piecewise constant, then we can simulate exact solutions. In the case where these parameters are time dependent, one can generate in a straightforward manner

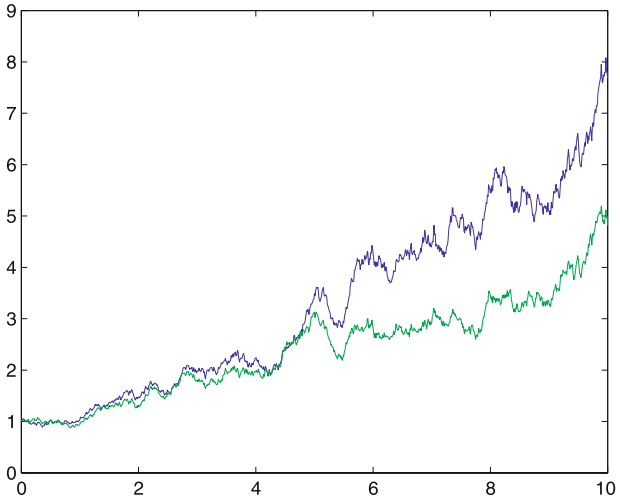

Fig. 2.4.1. Trajectory of a two-dimensional Black-Scholes model with parameters $S_0^1 = S_0^2 = 1$, $a_1 = a_2 = 0.1$, $b_1 = b_2 = 0.2$ and $\varrho = 0.8$

almost exact solutions by using, for instance, the trapezoidal rule. The main advantage of the multi-dimensional Black-Scholes model, which made it so popular, is that it has an explicit solution for the entire market dynamics and allows a wide range of easy calculations.

For a two-dimensional Black-Scholes model with $\boldsymbol{B}^1$ and $\boldsymbol{B}^2$ we obtain the following exact solution

$$S_t^1 = S_0^1 \exp\left\{\left(a_1 - \frac{1}{2}b_1^2\right)t + b_1 W_t^1\right\}, \tag{2.4.16}$$

$$S_t^2 = S_0^2 \exp\left\{\left(a_2 - \frac{1}{2}b_2^2\right)t + b_2(\varrho W_t^1 + \sqrt{1-\varrho^2}W_t^2)\right\}, \tag{2.4.17}$$

for $t \in [0, \infty)$. The two components of the trajectory of this two-dimensional model are illustrated in Fig. 2.4.1 for the parameter choice $S_0^1 = S_0^2 = 1$, $a_1 = a_2 = 0.1$, $b_1 = b_2 = 0.2$ and $\varrho = 0.8$.

Multi-dimensional Linear Diffusion Model

The continuous time limits of some popular time series models in finance can be described by a multi-dimensional ARCH diffusion model, see Sect. 1.7. The squared volatilities of this particular model emerge by assuming $\boldsymbol{\beta}^l$ to be zero in (2.4.5) for all $l \in \{1, 2, \ldots, m\}$.

We obtain an exact solution for the above multi-dimensional linear diffusion process using (2.4.6) in the following form

$$\boldsymbol{X}_t = \boldsymbol{\Psi}_t \left(\boldsymbol{X}_0 + \int_0^t \boldsymbol{\Psi}_s^{-1} \boldsymbol{\alpha}_s \, ds \right), \tag{2.4.18}$$

where, as before, $\boldsymbol{\Psi}_t$ satisfies (2.4.7). The multi-dimensional linear diffusion process can be simulated almost exactly provided that the matrices $\boldsymbol{A}$, $\boldsymbol{B}^1$, $\boldsymbol{B}^2, \ldots, \boldsymbol{B}^m$ commute. It simply remains to approximate the time integral in (2.4.18), which can be achieved with high accuracy for small time step size $\Delta > 0$ by using a quadrature formula, for instance, the trapezoidal rule when integrating between time discretization points. More precisely, we substitute for $t = t_i = i\Delta$, $i \in \{0, 1, \ldots\}$, the expression in (2.4.18) by the approximation

$$\boldsymbol{X}_{t_i}^{\Delta} = \boldsymbol{\Psi}_{t_i} \left(\boldsymbol{X}_0 + \sum_{k=0}^{i-1} \frac{\Delta}{2} \left(\boldsymbol{\Psi}_{t_{k+1}}^{-1} \boldsymbol{\alpha}_{t_{k+1}} + \boldsymbol{\Psi}_{t_k}^{-1} \boldsymbol{\alpha}_{t_k} \right) \right). \tag{2.4.19}$$

Other approximations are also possible. We remark that as a consequence of the methods described in Kloeden & Platen (1999) and those we will present later in this book, it is straightforward to show that the almost exact solution which uses the trapezoidal rule converges with strong order one in the sense of Kloeden & Platen (1999) and as we will define later in this book. By making the time step size sufficiently small any desired level of accuracy can be achieved. What is important in formula (2.4.19) is that errors do not propagate here, as can be the case with more general discrete-time approximations that will be discussed later in the book. In this sense, the almost exact solutions we simulate by the above method are still very accurate also over long periods of time. Numerical stability issues do not play any role in this context.

Let us now simulate a two-dimensional linear diffusion process with the following parameter matrices

$$\boldsymbol{A} = \begin{pmatrix} -\kappa_1 & 0 \\ 0 & -\kappa_2 \end{pmatrix}, \quad \boldsymbol{B}^1 = \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \varrho \end{pmatrix}, \quad \boldsymbol{B}^2 = \begin{pmatrix} 0 & 0 \\ 0 & \gamma_2 \sqrt{1 - \varrho^2} \end{pmatrix} \tag{2.4.20}$$

and vector

$$\boldsymbol{\alpha} = \begin{pmatrix} \kappa_1 \bar{x}_1 \\ \kappa_2 \bar{x}_2 \end{pmatrix}. \tag{2.4.21}$$

Then the vector SDE (2.4.5) simplifies to

$$dX_t^1 = \kappa_1 (\bar{x}_1 - X_t^1) dt + \gamma_1 X_t^1 \, dW_t^1, \tag{2.4.22}$$

$$dX_t^2 = \kappa_2 (\bar{x}_2 - X_t^2) dt + \gamma_2 X_t^2 \left(\varrho \, dW_t^1 + \sqrt{1 - \varrho^2} \, dW_t^2 \right), \tag{2.4.23}$$

for $t \in [0, \infty)$ with $X_0^1, X_0^2 > 0$. Note that the matrices $\boldsymbol{A}$, $\boldsymbol{B}^1$ and $\boldsymbol{B}^2$ commute, hence we can use the expression (2.4.9) for determining the fundamental matrix $\boldsymbol{\Psi}_t$. In the considered two-dimensional example the logarithm of this matrix equals

$$\ln(\boldsymbol{\Psi}_t) = \begin{pmatrix} \left(-\kappa_1 - \frac{1}{2}\gamma_1^2\right) t + \gamma_1 W_t^1 & 0 \\ 0 & \left(-\kappa_2 - \frac{1}{2}\gamma_2^2\right) t + \gamma_2 \left(\varrho W_t^1 + \sqrt{1 - \varrho^2} W_t^2 \right) \end{pmatrix}, \tag{2.4.24}$$

for $t_0 = 0$ and $t \in [0, \infty)$. Here the logarithm is interpreted elementwise. With this fundamental matrix the almost exact solution of (2.4.22) and (2.4.23) equals

$$X_{t_i}^{\Delta,1} = \exp\left\{\left(-\kappa_1 - \frac{1}{2}\gamma_1^2\right) t_i + \gamma_1 W_{t_i}^1\right\}$$
$$\times\left(X_0^1 + \kappa_1 \bar{x}_1 \sum_{k=0}^{i-1} \frac{\Delta}{2}\left[\exp\left\{\left(\kappa_1 + \frac{1}{2}\gamma_1^2\right) t_{k+1} - \gamma_1 W_{t_{k+1}}^1\right\}\right.\right.$$
$$\left.\left. + \exp\left\{\left(\kappa_1 + \frac{1}{2}\gamma_1^2\right) t_k - \gamma_1 W_{t_k}^1\right\}\right]\right), \tag{2.4.25}$$

$$X_{t_i}^{\Delta,2} = \exp\left\{\left(-\kappa_2 - \frac{1}{2}\gamma_2^2\right) t_i + \gamma_2(\varrho W_{t_i}^1 + \sqrt{1-\varrho^2} W_{t_i}^2)\right\}$$
$$\times\left(X_0^2 + \kappa_2 \bar{x}_2 \sum_{k=0}^{i-1} \frac{\Delta}{2}\left[\exp\left\{\left(\kappa_2 + \frac{1}{2}\gamma_2^2\right) t_{k+1} - \gamma_2(\varrho W_{t_{k+1}}^1 + \sqrt{1-\varrho^2}\, W_{t_{k+1}}^2)\right\}\right.\right.$$
$$\left.\left. + \exp\left\{\left(\kappa_2 + \frac{1}{2}\gamma_2^2\right) t_k - \gamma_2(\varrho W_{t_k}^1 + \sqrt{1-\varrho^2}\, W_{t_k}^2)\right\}\right]\right), \tag{2.4.26}$$

for $t_i = \Delta i, i \in \{0, 1, \ldots\}$. Given that we use a highly accurate numerical approximation for calculating the integrals in (2.4.25) and (2.4.26) we can simulate trajectories of X^1 and X^2 almost exactly, that means with any desired accuracy. This can be done by using the trapezoidal rule when approximating the time integrals in (2.4.18) and a sufficiently fine time discretization.

In Fig. 2.4.2 we display a trajectory of a two-dimensional linear diffusion process with the correlation parameter $\varrho = 0.8$, initial values $X_0^1 = X_0^2 = 1$ and $\bar{x}_1 = \bar{x}_2 = 0.5$, $\kappa_1 = \gamma_1 = 1$ and $\kappa_2 = \gamma_2 = 2$. One notes in this figure the similarities. On the other hand there are, of course differences in the paths of the two processes.

We refer to Kloeden & Platen (1999) for a list of other specific examples of explicitly solvable multi-dimensional SDEs. The above models can be further generalized by using some time changes and applying time changed Lévy processes. For instance, the above Black-Scholes model can be generalized by subordination to multi-dimensional SDEs driven by Lévy processes, yielding exponential Lévy market models, see Barndorff-Nielsen & Shephard (2001), Geman et al. (2001) and Eberlein (2002).

2.5 Almost Exact Solutions by Conditioning

For some multi-dimensional SDEs it is possible to simulate one component using its marginal distribution or other exact or almost exact methods. The

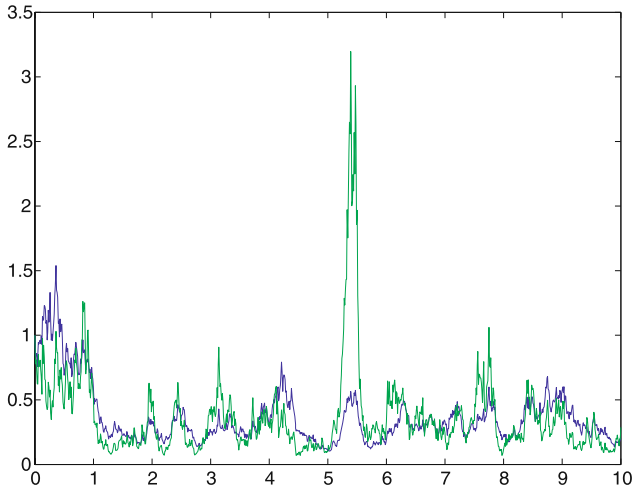

Fig. 2.4.2. Two correlated linear diffusion processes obtained by setting $\varrho = 0.8$, $X_0^1 = X_0^2 = 1$, $\bar{x}_1 = \bar{x}_2 = 0.5$, $\kappa_1 = \gamma_1 = 1$ and $\kappa_2 = \gamma_2 = 2$

second component can be conditioned on the first one, and then simulated exactly or almost exactly. This can eventually be continued to further components. Let us illustrate and describe this concept by applying it to the multi-dimensional Heston model, see Heston (1993) and Sect. 1.6.

One-dimensional Heston Model

The Heston model is probably the most popular stochastic volatility model used in finance. The reason why it is used so widely in practice is that Heston (1993) was able to derive an exact formula for European call and put option prices under this model. This underlines the relevance of the current chapter with its focus on exact solutions.

The Heston model can be expressed by a two-dimensional SDE. Its first component models the stock price S_t while its second component characterizes the squared volatility V_t of the underlying stock. This system of SDEs is typically written in the form

$$dS_t = rS_t dt + \sqrt{V_t} S_t \left[\varrho dW_t^1 + \sqrt{1-\varrho^2} dW_t^2 \right], \tag{2.5.1}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t^1, \tag{2.5.2}$$

for $t \in [0, \infty)$. Here W^1 and W^2 are two independent Wiener processes, and ϱ represents the correlation parameter. Furthermore, r denotes the short rate,

κ the speed of adjustment, θ the average squared volatility and σ the volatility of the squared volatility.

There is a wide literature on approximating the Heston model by simulation. Recent studies include Broadie & Kaya (2006), Smith (2007) and Andersen (2008) who discuss exact or almost exact simulation techniques for the Heston model. For the purpose of simulation it is convenient to employ the logarithmic transformation $X_t = \ln(S_t)$ instead of using (2.5.1) directly. Therefore, by the Itô formula we obtain the SDE

$$\begin{aligned} dX_t &= \left(r - \frac{1}{2}V_t\right) dt + \sqrt{V_t}\left[\varrho\, dW_t^1 + \sqrt{1-\varrho^2}\, dW_t^2\right], \\ dV_t &= \kappa(\theta - V_t)\, dt + \sigma\sqrt{V_t}\, dW_t^1, \end{aligned} \tag{2.5.3}$$

for $t \in [0, \infty)$.

In Broadie & Kaya (2006), given V_{t_i}, $i \in \{0, 1, \ldots\}$, the value $V_{t_i+\Delta}$ is sampled directly from the non-central chi-square distribution with δ degrees of freedom and non-centrality parameter

$$\lambda = \frac{4\kappa e^{-\kappa\Delta}}{\sigma^2(1 - e^{-\kappa\Delta})} V_{t_i}.$$

That is

$$V_{t_i+\Delta} = \frac{\sigma^2(1 - e^{-\kappa\Delta})}{4\kappa} \chi^2_{(\delta,\lambda)}, \tag{2.5.4}$$

where $\delta = \frac{4\theta\kappa}{\sigma^2}$. Here $\chi^2(\delta, \lambda)$ is sampled from the non-central chi-square distribution function

$$F_X(x) = \sum_{k=0}^{\infty} \frac{\exp\left\{-\frac{\lambda}{2}\right\}\left(\frac{\lambda}{2}\right)^k}{k\,!} \left(1 - \frac{\Gamma\left(\frac{x}{2}; \frac{\delta+2k}{2}\right)}{\Gamma\left(\frac{\delta+2k}{2}\right)}\right), \tag{2.5.5}$$

where

$$\Gamma(u; a) = \int_u^{\infty} t^{a-1} \exp\{-t\}\, dt$$

is the incomplete gamma function for $u \geq 0$, $a > -1$. Note that the non-central chi-square distribution arises as a scalar version of the non-central Wishart distribution. Sampling from the non-central chi-square distribution is discussed, for instance, in Glasserman (2004). The resulting simulation method for V is exact.

In order to obtain an exact scheme for the simulation of the asset price process in the Heston model we note that for $t_i = \Delta i, i \in \{0, 1, \ldots\}$, we obtain

$$V_{t_{i+1}} = V_{t_i} + \int_{t_i}^{t_{i+1}} \kappa(\theta - V_u) du + \sigma \int_{t_i}^{t_{i+1}} \sqrt{V_u} dW_u^1. \tag{2.5.6}$$

Hence,

$$\int_{t_i}^{t_{i+1}} \sqrt{V_u} dW_u^1 = \sigma^{-1} \left(V_{t_{i+1}} - V_{t_i} - \kappa\theta\Delta + \kappa \int_{t_i}^{t_{i+1}} V_u du \right). \tag{2.5.7}$$

Additionally,

$$X_{t_{i+1}} = X_{t_i} + \int_{t_i}^{t_{i+1}} \left(r - \frac{1}{2} V_u \right) du + \varrho \int_{t_i}^{t_{i+1}} \sqrt{V_u} dW_u^1 + \sqrt{1-\varrho^2} \int_{t_i}^{t_{i+1}} \sqrt{V_u} dW_u^2. \tag{2.5.8}$$

Substituting (2.5.7) into (2.5.8) we obtain

$$\begin{aligned} X_{t_{i+1}} = X_{t_i} + r\Delta + \frac{\varrho}{\sigma}(V_{t_{i+1}} - V_{t_i} - \kappa\theta\Delta) + \left(\frac{\varrho\kappa}{\sigma} - \frac{1}{2} \right) \int_{t_i}^{t_{i+1}} V_u \, du \\ + \sqrt{1-\varrho^2} \int_{t_i}^{t_{i+1}} \sqrt{V_u} \, dW_u^2. \end{aligned} \tag{2.5.9}$$

Furthermore, the distribution of $\int_{t_i}^{t_{i+1}} \sqrt{V_u} dW_u^2$, given the path generated by V_t, is conditionally normal with mean zero and variance $\int_{t_i}^{t_{i+1}} V_u du$, because V is independent of the Brownian motion W^2, that is

$$\int_{t_i}^{t_{i+1}} \sqrt{V_u} dW_u^2 \sim N\left(0, \int_{t_i}^{t_{i+1}} V_u du\right). \tag{2.5.10}$$

Broadie & Kaya (2006) discuss in detail how to obtain explicitly $\int_{t_i}^{t_{i+1}} V_u du$. However, it is also possible to approximate $\int_{t_i}^{t_{i+1}} V_u du$ given the path of the process V. This approximation can be achieved with high accuracy by a quadrature formula such as the trapezoidal rule, as previously discussed, which results in an efficient almost exact simulation technique by conditioning for the Heston model.

Multi-dimensional Heston Model with Independent Prices

Let us now define a particular version of the multi-dimensional Heston model given via the following SDE

$$d\boldsymbol{S}_t = \boldsymbol{A}_t \left(\boldsymbol{r} dt + \sqrt{\boldsymbol{B}_t} d\boldsymbol{W}_t \right). \tag{2.5.11}$$

Here $\boldsymbol{S} = \{\boldsymbol{S}_t = (S_t^1, S_t^2, \ldots, S_t^d)^\top, t \in [0, \infty)\}$ is a vector process and $\boldsymbol{A}_t = [A_t^{i,j}]_{i,j=1}^d$ is a diagonal matrix process with elements

$$A_t^{i,j} = \begin{cases} S_t^i & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.5.12}$$

Additionally, $\boldsymbol{r} = (r_1, r_2, \ldots, r_d)^\top$ is a d-dimensional vector and $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^d)^\top, t \in [0, \infty)\}$ is a d-dimensional vector of correlated Wiener processes. Moreover, $\boldsymbol{B}_t = [B_t^{i,j}]_{i,j=1}^d$ is a matrix process with elements

$$B_t^{i,j} = \begin{cases} \Sigma_t^{i,i} & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.5.13}$$

Note that $\boldsymbol{B}$ is the generalization of V in the one-dimensional case. Here, the matrix process $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_t = \Sigma_t^{i,j}, t \geq 0\}$ is a matrix SR-process given by the SDE (2.3.66). Therefore, $\boldsymbol{B}_t$ can be constructed from the diagonal elements of $\boldsymbol{\Sigma}_t$. Recall, that these elements $\Sigma_t^{1,1}, \Sigma_t^{2,2}, \ldots, \Sigma_t^{d,d}$ form correlated SR-processes. For simplicity, let us first assume that $\boldsymbol{B}$ is independent of $\boldsymbol{W}$.

The simulation of such a multi-dimensional version of the Heston model is straightforward and simply generalizes the simulation of the one-dimensional Heston model. Let us illustrate this on a two-dimensional example. The corresponding two-dimensional SDE for the two asset prices can be represented as

$$\begin{aligned} dS_t^1 &= S_t^1 r_1 dt + \sqrt{\Sigma_t^{1,1}} dW_t^1 \\ dS_t^2 &= S_t^2 r_2 dt + \sqrt{\Sigma_t^{2,2}} \left[\varrho dW_t^1 + \sqrt{1-\varrho^2} dW_t^2 \right], \end{aligned} \tag{2.5.14}$$

for $t \in [0, \infty)$. Here $\Sigma^{1,1}$ and $\Sigma^{2,2}$ are diagonal elements of the 2×2 matrix square root process given by the SDE (2.3.66). Note that $\Sigma^{1,1}$ and $\Sigma^{2,2}$ are always positive, hence $\sqrt{\Sigma^{1,1}}$ and $\sqrt{\Sigma^{2,2}}$ are well defined. Furthermore, the logarithmic transformation $\boldsymbol{X}_t = \ln(\boldsymbol{S}_t)$ yields the following SDE

$$\begin{aligned} dX_t^1 &= \left(r_1 - \frac{1}{2}\Sigma_t^{1,1} \right) dt + \sqrt{\Sigma_t^{1,1}} dW_t^1 \\ dX_t^2 &= \left(r_2 - \frac{1}{2}\Sigma_t^{2,2} \right) dt + \sqrt{\Sigma_t^{2,2}} \left[\varrho dW_t^1 + \sqrt{1-\varrho^2} dW_t^2 \right], \end{aligned} \tag{2.5.15}$$

for $t \in [0, \infty)$. This results in the following representations

$$\begin{aligned} X_{t_{i+1}}^1 &= X_{t_i}^1 + r_1 \Delta - \frac{1}{2} \int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du + \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,1}} dW_u^1 \\ X_{t_{i+1}}^2 &= X_{t_i}^2 + r_2 \Delta - \frac{1}{2} \int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du + \varrho \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}} dW_u^1 \\ &\quad + \sqrt{1-\varrho^2} \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}} dW_u^2, \end{aligned} \tag{2.5.16}$$

where

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,1}} dW_u^1 \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du\right) \tag{2.5.17}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}} dW_u^1 \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du\right) \tag{2.5.18}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}} dW_u^2 \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du\right). \tag{2.5.19}$$

Given that we can approximate almost exactly $\int_{t_i}^{t_{i+t}} \Sigma_u^{1,1} du$ and $\int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du$ it is possible to simulate this two-dimensional Heston model very accurately. We have discussed in Sect. 2.2 and 2.3 how to simulate the matrix SR-process $\boldsymbol{\Sigma}$, which applies here. The terms in (2.5.17) and (2.5.18) can be generated as corresponding Gaussian random variables.

Multi-dimensional Heston Model with Correlated Prices

Let us now consider another multi-dimensional version of the Heston model, which allows correlation of the volatility vector $\boldsymbol{\Sigma}$ with the vector asset price process $\boldsymbol{S}$. We define this generalization by the system of SDEs

$$d\boldsymbol{S}_t = \boldsymbol{A}_t \left[\boldsymbol{r}\, dt + \sqrt{\boldsymbol{B}_t} \left(\boldsymbol{C}\, d\boldsymbol{W}_t^1 + \boldsymbol{D}\, d\boldsymbol{W}_t^2 \right) \right] \tag{2.5.20}$$

$$d\boldsymbol{\Sigma}_t = (\boldsymbol{a} - \boldsymbol{E}\boldsymbol{\Sigma}_t)\, dt + \boldsymbol{F}\sqrt{\boldsymbol{B}_t}\, d\boldsymbol{W}_t^1, \tag{2.5.21}$$

for $t \in [0, \infty)$. Here, $\boldsymbol{S} = \{\boldsymbol{S}_t = (S_t^1, S_t^2, \ldots, S_t^d)^\top, t \in [0, \infty)\}$ and $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)^\top$. $\boldsymbol{A}_t = [A_t^{i,j}]_{i,j=1}^d$ is a matrix with elements as in (2.5.12) and $\boldsymbol{B}_t = [B_t^{i,j}]_{i,j=1}^d$ is a matrix with elements as in (2.5.13). Additionally, $\boldsymbol{C} = [C^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$C^{i,j} = \begin{cases} \varrho_i & \text{for } i = j \\ 0 & \text{otherwise,} \end{cases} \tag{2.5.22}$$

and $\boldsymbol{D} = [D^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$D^{i,j} = \begin{cases} \sqrt{1 - \varrho_i^2} & \text{for } i = j \\ 0 & \text{otherwise,} \end{cases} \tag{2.5.23}$$

where $\varrho \in [-1, 1]$, $i \in \{1, 2, \ldots, d\}$. Moreover, $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_t = (\Sigma_t^{1,1}, \Sigma_t^{2,2}, \ldots, \Sigma_t^{d,d})^\top, t \in [0, \infty)\}$ and $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^\top$. $\boldsymbol{E} = [E^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$E^{i,j} = \begin{cases} b_i & \text{for } i = j \\ 0 & \text{otherwise,} \end{cases} \tag{2.5.24}$$

and $\boldsymbol{F} = [F^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$F^{i,j} = \begin{cases} \sigma_i & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.5.25}$$

Furthermore, $\boldsymbol{W}^1 = \{\boldsymbol{W}_t^1 = (W_t^{1,1}, W_t^{1,2}, \ldots, W_t^{1,d})^\top, t \in [0, \infty)\}$ and $\boldsymbol{W}^2 = \{\boldsymbol{W}_t^2 = (W_t^{2,1}, W_t^{2,2}, \ldots, W_t^{2,d})^\top, t \in [0, \infty)\}$ are vectors of correlated Wiener processes which are independent of each other.

Let us now illustrate the simulation of a two-dimensional Heston model of this type. The two-dimensional squared volatility process can be described by the following SDE

$$d\Sigma_t^{1,1} = \left(a_1 - b_1\Sigma_t^{1,1}\right) dt + \sigma_1\sqrt{\Sigma_t^{1,1}}\, dW_t^{1,1}, \tag{2.5.26}$$

$$d\Sigma_t^{2,2} = \left(a_2 - b_2\Sigma_t^{2,2}\right) dt + \sigma_2\sqrt{\Sigma_t^{2,2}}\, dW_t^{1,2}, \tag{2.5.27}$$

for $t \in [0, \infty)$. Moreover, the 2-dimensional asset price process is given by

$$dS_t^1 = S_t^1 r_1 dt + S_t^1\sqrt{\Sigma_t^{1,1}}\left(\varrho_1\, dW_t^{1,1} + \sqrt{1-\varrho_1^2}\, dW_t^{2,1}\right), \tag{2.5.28}$$

$$dS_t^2 = S_t^2 r_2 dt + S_t^2\sqrt{\Sigma_t^{2,2}}\left(\varrho_2\, dW_t^{1,2} + \sqrt{1-\varrho_2^2}\, dW_t^{2,2}\right), \tag{2.5.29}$$

for $t \in [0, \infty)$. Here the elements of the $\boldsymbol{\Sigma}$ process can be simulated as diagonal elements of a matrix SR-process as described in Sects. 2.2 and 2.3. Given these paths we may simulate the path of $\boldsymbol{X}_t = \ln(\boldsymbol{S}_t)$ similar as in the algorithm (2.5.9). That is

$$X_{t_{i+1}}^1 = X_{t_i}^1 + r_1\Delta + \frac{\varrho_1}{\sigma_1}\left(\Sigma_{t_{i+1}}^{1,1} - \Sigma_{t_i}^{1,1} - a_1\Delta\right) + \left(\frac{\varrho_1 b_1}{\sigma_1} - \frac{1}{2}\right)\int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du$$
$$+ \sqrt{1-\varrho_1^2}\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,1}}\, dW_u^{2,1},$$

$$X_{t_{i+1}}^2 = X_{t_i}^2 + r_2\Delta + \frac{\varrho_2}{\sigma_2}\left(\Sigma_{t_{i+1}}^{2,2} - \Sigma_{t_i}^{2,2} - a_2\Delta\right) + \left(\frac{\varrho_2 b_2}{\sigma_2} - \frac{1}{2}\right)\int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du$$
$$+ \sqrt{1-\varrho_2^2}\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}}\, dW_u^{2,2}.$$

Here,

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,1}} dW_u^{2,1} \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du\right), \tag{2.5.30}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}} dW_u^{2,2} \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du\right) \tag{2.5.31}$$

have now to be simulated as two correlated Gaussian random variables with zero mean and variances $\int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du$ and $\int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du$. The correlation parameters are $\varrho_1, \varrho_2 \in [-1, 1]$. This simulation method is almost exact, since the only approximation used here is the approximation of the time integrals $\int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du$ and $\int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du$. Numerical errors are not propagated.

Another Heston Model

There exist alternative ways to construct multi-dimensional Heston type models. Two possible generalizations are discussed in Gouriéroux & Sufana (2004)

and Da Fonseca, Grasselli & Tebaldi (2008). For instance, Gouriéroux & Sufana (2004) consider a d-dimensional stochastic volatility model of the form

$$d\boldsymbol{S}_t = \boldsymbol{A}_t \left(r\mathbf{1}\, dt + \sqrt{\boldsymbol{\Sigma}_t}\, d\boldsymbol{W}_t^2 \right), \tag{2.5.32}$$
$$d\boldsymbol{\Sigma}_t = \left(\boldsymbol{\Omega}\boldsymbol{\Omega}^\top + \boldsymbol{M}\boldsymbol{\Sigma}_t + \boldsymbol{\Sigma}_t\boldsymbol{M}^\top \right) dt + \sqrt{\boldsymbol{\Sigma}_t}\, d\boldsymbol{W}_t^1 \boldsymbol{Q} + \boldsymbol{Q}^\top d(\boldsymbol{W}_t^1)^\top \sqrt{\boldsymbol{\Sigma}_t},$$

for $t \in [0, \infty)$. Here, $\boldsymbol{S} = \{\boldsymbol{S}_t = (S_t^1, S_t^2, \ldots, S_t^d)^\top, t \in [0, \infty)\}$ is a vector asset price process and $\boldsymbol{\Sigma}_t = [\Sigma_t^{i,j}]_{i,j=1}^d$ is a matrix squared volatility process. Moreover, elements of the matrix $\boldsymbol{A}_t = [A_t^{i,j}]_{i,j=1}^d$ are defined by (2.5.12) and $\mathbf{1} = (1, 1, \ldots, 1)^\top$. Here, $\boldsymbol{W}^2 = \{\boldsymbol{W}_t^2 = (W_t^{2,1}, W_t^{2,2}, \ldots, W_t^{2,d})^\top, t \in [0, \infty)\}$ is a vector of independent Wiener processes. In the second equation $\boldsymbol{\Omega}$, $\boldsymbol{M}$ and $\boldsymbol{Q}$ are $d \times d$ parameter matrices with $\boldsymbol{\Omega}$ being an invertible matrix. In order to preserve strict positivity and the mean reverting feature of the volatility, the matrix $\boldsymbol{M}$ is assumed to be negative semi-definite, while $\boldsymbol{\Omega}$ satisfies

$$\boldsymbol{\Omega}\boldsymbol{\Omega}^\top = \beta \boldsymbol{\Omega}^\top \boldsymbol{\Omega} \tag{2.5.33}$$

with real valued parameter $\beta = d - 1$, see Bru (1991). Additionally, $\boldsymbol{W}_t^1 = [W_t^{1,i,j}]_{i,j=1}^d$ is a matrix Wiener process independent of $\boldsymbol{W}_t^2$.

In the two-dimensional case this model is represented by the SDEs

$$dS_t^1 = rS_t^1\, dt + S_t^1 \sqrt{\Sigma_t^{1,1}}\, dW_t^{2,1} + S_t^1 \sqrt{\Sigma_t^{1,2}}\, dW_t^{2,2}, \tag{2.5.34}$$
$$dS_t^2 = rS_t^2\, dt + S_t^2 \sqrt{\Sigma_t^{2,1}}\, dW_t^{2,1} + S_t^2 \sqrt{\Sigma_t^{2,2}}\, dW_t^{2,2}, \tag{2.5.35}$$

for $t \in [0, \infty)$. The SDE for $\boldsymbol{X}_t = \ln(\boldsymbol{S}_t)$ is given by

$$dX_t^1 = \left(r - \frac{1}{2}\Sigma_t^{1,1} - \frac{1}{2}\Sigma_t^{1,2} \right) dt + \sqrt{\Sigma_t^{1,1}} dW_t^{2,1} + \sqrt{\Sigma_t^{1,2}} dW_t^{2,2}, \tag{2.5.36}$$
$$dX_t^2 = \left(r - \frac{1}{2}\Sigma_t^{2,1} - \frac{1}{2}\Sigma_t^{1,2} \right) dt + \sqrt{\Sigma_t^{2,1}} dW_t^{2,1} + \sqrt{\Sigma_t^{2,2}} dW_t^{2,2}, \tag{2.5.37}$$

for $t \in [0, \infty)$. Therefore, since $\boldsymbol{\Sigma}$ is independent of $\boldsymbol{W}^2$ we obtain an almost exact simulation algorithm for $\boldsymbol{X}$ constructed as

$$\begin{aligned} X_{t_{i+1}}^1 &= X_{t_i}^1 + r\Delta - \frac{1}{2}\int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du - \frac{1}{2}\int_{t_i}^{t_{i+1}} \Sigma_u^{1,2} du + \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,1}}\, dW_u^{2,1} \\ &\quad + \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,2}}\, dW_u^{2,2}, \\ X_{t_{i+1}}^2 &= X_{t_i}^2 - \frac{1}{2}\int_{t_i}^{t_{i+1}} \Sigma_u^{2,1} du - \frac{1}{2}\int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du + \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,1}}\, dW_u^{2,1} \\ &\quad + \int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}}\, dW_u^{2,2}. \end{aligned}$$

Here,

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,1}} dW_u^{2,1} \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{1,1} du\right) \tag{2.5.38}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{1,2}} dW_u^{2,2} \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{1,2} du\right) \tag{2.5.39}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,1}} dW_u^{2,1} \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{2,1} du\right) \tag{2.5.40}$$

$$\int_{t_i}^{t_{i+1}} \sqrt{\Sigma_u^{2,2}} dW_u^{2,2} \sim N\left(0, \int_{t_i}^{t_{i+1}} \Sigma_u^{2,2} du\right) \tag{2.5.41}$$

are independent Gaussian random variables with zero means and variance as indicated. Given that we can approximate $\int_{t_i}^{t_{i+1}} \Sigma_u^{k,j} du$ for $k, j \in \{1, 2\}$ accurately via the trapezoidal rule this simulation is again almost exact.

2.6 Almost Exact Simulation by Time Change

Another useful technique, which can be applied in the simulation of multi-dimensional SDEs, is the almost exact simulation by *stochastic time change*. This method can be used, for instance, for the simulation of solutions of SDEs which can be written as time changed solutions of SDEs that permit almost exact simulation. The time change used can be stochastic and needs to be simulated exactly or almost exactly. We will illustrate this technique by the simulation of two models, a multi-dimensional ARCH diffusion model and a multi-dimensional generalized minimal market model (MMM).

One-dimensional ARCH Diffusion Model

An interesting stochastic volatility model is the ARCH diffusion model, which is represented by a two-dimensional SDE, see (1.7.32)–(1.7.33). Again the first SDE models the stock price while the second one the squared volatility of the underlying stock. The only difference between the Heston model and the ARCH diffusion model is in the specification of the diffusion coefficient function of the squared volatility process, which is multiplicative. That is,

$$dS_t = rS_t\, dt + \sqrt{V_t} S_t \left[\varrho dW_t^1 + \sqrt{1-\varrho^2}\, dW_t^2\right], \tag{2.6.1}$$

$$dV_t = \kappa(\theta - V_t)\, dt + \sigma V_t\, dW_t^1, \tag{2.6.2}$$

for $t \in [0, \infty)$. Here W^1 and W^2 are two independent Wiener processes, and ϱ represents the correlation between the noises driving the return process and the volatility process. In the given case we can simulate the squared volatility process V almost exactly by approximating again some time integral via the trapezoidal rule. Here the exact representation is of the form

$$V_{t_i} = \exp\left\{\left(-\kappa - \frac{1}{2}\sigma^2\right) t_i + \sigma W^1_{t_i}\right\}$$
$$\times \left(V_{t_0} + \kappa\,\theta \sum_{k=0}^{i-1} \int_{t_k}^{t_{k+1}} \exp\left\{\left(\kappa + \frac{1}{2}\sigma^2\right) s - \sigma W^1_s\right\} ds\right), \quad (2.6.3)$$

which leads to the recursive almost exact approximation

$$V^{\Delta}_{t_i} = \exp\left\{\left(-\kappa - \frac{1}{2}\sigma^2\right) t_i + \sigma W^1_{t_i}\right\}$$
$$\times \left(V_{t_0} + \kappa\,\theta \sum_{k=0}^{i-1} \frac{\Delta}{2}\left[\exp\left\{\left(\kappa + \frac{1}{2}\sigma^2\right) t_k - \sigma W^1_{t_k}\right\}\right.\right.$$
$$\left.\left. + \exp\left\{\left(\kappa + \frac{1}{2}\sigma^2\right) t_{k+1} - \sigma W^1_{t_{k+1}}\right\}\right]\right) \quad (2.6.4)$$

for $t_i = \Delta i, i \in \{0, 1, \ldots\}$. Note that this simulation requires sampling from a Gaussian law in contrast to the simulation of the squared volatility process in the Heston model, where we sampled from the non-central chi-square distribution. It is now straightforward to correlate the squared volatility process V with the asset price process S.

In order to simulate the asset price process S we consider its discounted version $\bar{S}_t = \frac{S_t}{S^0_t}$, where $S^0_t = \exp\{rt\}$ is the savings account, for $t \in [0, \infty)$. This results in the following SDE

$$d\bar{S}_t = \sqrt{V_t}\,\bar{S}_t \left[\varrho\, dW^1_t + \sqrt{1-\varrho^2}\, dW^2_t\right], \quad (2.6.5)$$

for $t \in [0, \infty)$. Additionally, we obtain by the Itô formula

$$d\ln(\bar{S}_t) = -\frac{1}{2} V_t\, dt + \sqrt{V_t}\left[\varrho\, dW^1_t + \sqrt{1-\varrho^2}\, dW^2_t\right], \quad (2.6.6)$$

for $t \in [0, \infty)$. Equation (2.6.6) for $\ln(\bar{S}_t) = \ln(\bar{S}(\tau(t)))$ turns out to be that of a time changed, drifted Wiener process $\ln(\bar{S}) = \{\ln(\bar{S}(\tau(t))), t \in [0, \infty)\}$ with SDE of the form

$$d\ln(\bar{S}(\tau(t))) = -\frac{1}{2}\, d\tau(t) + \left[\varrho\, d\tilde{W}^1_{\tau(t)} + \sqrt{1-\varrho^2}\, d\tilde{W}^2_{\tau(t)}\right], \quad (2.6.7)$$

for $t \in [0, \infty)$. Note that we make a difference between $\bar{S}_\cdot$ and $\bar{S}(\cdot)$. Here the time change is stochastic and can be obtained by approximation of the following time integral

$$\tau(t) = \tau(0) + \int_0^t V_s ds, \quad (2.6.8)$$

for $t \in [0, \infty)$. Furthermore, in τ-time the processes $\tilde{W}^1$ and $\tilde{W}^2$ are independent Wiener processes. Under the above ARCH diffusion model the asset price is then obtained as $S_t = S^0_t\, \bar{S}(\tau(t)) = S^0_t \bar{S}_t$.

Multi-dimensional ARCH Diffusion

Let us now define a multi-dimensional ARCH diffusion model as follows

$$d\boldsymbol{S}_t = \boldsymbol{A}_t \left[\boldsymbol{r}\, dt + \sqrt{\boldsymbol{B}_t} \left(\boldsymbol{C}\, d\boldsymbol{W}_t^1 + \boldsymbol{D}\, d\boldsymbol{W}_t^2 \right) \right] \tag{2.6.9}$$

$$d\boldsymbol{V}_t = (\boldsymbol{a} - \boldsymbol{E}\boldsymbol{V}_t)\, dt + \boldsymbol{F}\boldsymbol{B}_t\, d\boldsymbol{W}_t^1, \tag{2.6.10}$$

for $t \in [0, \infty)$. Here, $\boldsymbol{S} = \{\boldsymbol{S}_t = (S_t^1, S_t^2, \ldots, S_t^d)^\top, t \in [0, \infty)\}$, $\boldsymbol{A}_t = [A_t^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements as in (2.5.12), $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)^\top$. Furthermore, $\boldsymbol{B}_t = [B_t^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$B_t^{i,j} = \begin{cases} V_t^i & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.6.11}$$

Additionally, $\boldsymbol{C} = [C^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements as in (2.5.22) and $\boldsymbol{D} = [D^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements as in (2.5.23).

Moreover, $\boldsymbol{V} = \{\boldsymbol{V}_t = (V_t^1, V_t^2, \ldots, V_t^d)^\top, t \in [0, \infty)\}$, $\boldsymbol{a} = (a_1, a_2, \ldots, a_d)^\top$ and $\boldsymbol{E} = [E^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$E^{i,j} = \begin{cases} \kappa_i & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.6.12}$$

Furthermore, $\boldsymbol{F} = [F^{i,j}]_{i,j=1}^d$ is a diagonal matrix with elements

$$F^{i,j} = \begin{cases} \sigma_i & \text{for } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.6.13}$$

Finally, $\boldsymbol{W}^1 = \{\boldsymbol{W}_t^1 = (W_t^{1,1}, W_t^{1,2}, \ldots, W_t^{1,d})^\top, t \in [0, \infty)\}$ and $\boldsymbol{W}^2 = \{\boldsymbol{W}_t^2 = (W_t^{2,1}, W_t^{2,2}, \ldots, W_t^{2,d})^\top, t \in [0, \infty)\}$ are vectors of correlated Wiener processes but independent of each other. This setup then yields, with the previous methodology, almost exact approximate solutions for the multi-dimensional ARCH diffusion given in (2.6.9)–(2.6.10).

One-dimensional Stylized MMM

Let us now describe the minimal market model (MMM), introduced in Platen (2001), see (1.6.9)–(1.6.10). It has been designed for modeling the long term dynamics of the growth optimal portfolio (GOP), see Platen & Heath (2006), when denominated in a given currency. Under this model the discounted GOP, denoted by $\bar{S}^{\delta_*}$, forms a time transformed squared Bessel process of dimension four. This model reflects well various empirical properties of the long term dynamics of a diversified world stock index. In its stylized version it can be described by the SDE

$$d\bar{S}_t^{\delta_*} = \alpha_t^{\delta_*}\, dt + \sqrt{\bar{S}_t^{\delta_*} \alpha_t^{\delta_*}}\, dW_t, \tag{2.6.14}$$

for $t \in [0, \infty)$, with $\alpha_t^{\delta_*} = \alpha_0 \exp\{\eta t\}$, for initial scaling parameter $\alpha_0 > 0$ and net growth rate $\eta > 0$. This is equivalent to writing

$$d\bar{S}_t^{\delta_*} = 4d\varphi(t) + 2\sqrt{\bar{S}_t^{\delta_*}}\, d\tilde{W}_{\varphi(t)}, \tag{2.6.15}$$

for $t \in [0, \infty)$, which is the SDE of a squared Bessel process $X = \{X_{\varphi(t)} = \bar{S}_t^{\delta_*}, t \geq 0\}$ of dimension $\delta = 4$ in φ-time. Here, one has

$$\varphi(t) = \frac{\alpha_0}{4\eta} \left(\exp\{\eta t\} - 1 \right), \tag{2.6.16}$$

and the process $\tilde{W} = \{\tilde{W}_\varphi,\, \varphi \geq 0\}$ is a standard Wiener process in φ-time.

The normalized GOP under the stylized MMM is given by the ratio

$$Y_t = \frac{\bar{S}_t^{\delta_*}}{\alpha_t^{\delta_*}}. \tag{2.6.17}$$

By the Itô formula it follows that the process $Y = \{Y_t, t \in [0, T]\}$ is a square root process of dimension four with SDE

$$dY_t = (1 - \eta Y_t)\, dt + \sqrt{Y_t}\, dW_t \tag{2.6.18}$$

for $t \in \Re^+$ and initial value $Y_0 = \bar{S}_0^{\delta_*}/\alpha_0^{\delta_*}$. Note also that the quadratic variation of the square root of Y_t equals

$$[\sqrt{Y}]_t = \frac{1}{4} t, \tag{2.6.19}$$

for $t \in [0, \infty)$. Additionally, the volatility under the stylized MMM is of the form

$$\sigma_t = \frac{1}{\sqrt{Y_t}}, \tag{2.6.20}$$

for $t \in [0, \infty)$.

Hence, the value $\bar{S}^{\delta_*}$ at time t of the discounted GOP can be expressed in the form

$$\bar{S}_t^{\delta_*} = Y_t \alpha_t^{\delta_*} \tag{2.6.21}$$

for $t \in [0, \infty)$. Therefore, the GOP when expressed in units of the domestic currency can be represented by the product

$$S_t^{\delta_*} = S_t^0 \bar{S}_t^{\delta_*} = S_t^0 Y_t \alpha_t^{\delta_*} \tag{2.6.22}$$

for $t \in [0, \infty)$, where $S_t^0 = \exp\{\int_0^t r_s ds\}$ is the domestic savings account with $r = \{r_t, t \in [0, \infty)\}$ denoting the short rate process.

Multi-dimensional Stylized MMM

Let us now consider the MMM for a market with $d+1$ currencies, similar to the one described in Platen (2001) and Heath & Platen (2005). We can describe the value of the GOP in the ith currency denomination according to (2.6.22) by the expression

$$S_i^{\delta_*}(t) = \alpha_t^i Y_t^i S_i^i(t), \tag{2.6.23}$$

where $\alpha_t^i = \alpha_0^i \exp\{\eta^i t\}$ and $S_i^i(t) = \exp\{r^i t\}$, $i \in \{0, 1, \ldots, d\}$. Here η^i is the ith net growth rate and r^i is the short rate for the ith currency for $i \in \{0, 1, \ldots, d\}$. The ith normalized GOP satisfies then the SDE

$$dY_t^i = (1 - \eta^i Y_t^i)\, dt + \sqrt{Y_t^i}\, dW_t^i \tag{2.6.24}$$

for $t \in [0, \infty)$, with $Y_0^i > 0$. Here $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^{d+1})^\top, t \in [0, \infty)\}$ is a vector of correlated Wiener processes.

Let us now illustrate the simulation of a two-currency market based on the stylized MMM. We start with the simulation of two time changed 4×2 matrix Wiener processes $\boldsymbol{W}^1_{\varphi_1(t)} = [W^{1,i,j}_{\varphi_1(t)}]^{4,2}_{i,j=1}$ and $\boldsymbol{W}^2_{\varphi_2(t)} = [W^{2,i,j}_{\varphi_2(t)}]^{4,2}_{i,j=1}$, resulting from the same set of Gaussian random numbers. The elements of such a matrix are as in (2.3.25). The time changes are here

$$\varphi_i(t) = \frac{\alpha_0^i}{4\eta^i} \left(\exp\{\eta^i t\} - 1\right), \tag{2.6.25}$$

for $i \in \{1, 2\}$. In the next step we construct two 4×2 matrix Wiener processes $\tilde{\boldsymbol{W}}^1_{\varphi_1(t)} = \mathbf{I}_{4\times 4} \boldsymbol{W}^1_{\varphi_1(t)} \boldsymbol{\Sigma}^\top_{2\times 2}$ and $\tilde{\boldsymbol{W}}^2_{\varphi_2(t)} = \mathbf{I}_{4\times 4} \boldsymbol{W}^2_{\varphi_2(t)} \boldsymbol{\Sigma}^\top_{2\times 2}$, for $t \in [0, \infty)$ where the elements of $\boldsymbol{\Sigma}$ are as in (2.3.13). As a result of this transformation we obtain the 4×2 matrix processes $\tilde{\boldsymbol{W}}^1$ and $\tilde{\boldsymbol{W}}^2$ with correlated columns and independent rows. Additionally, the 4×2 matrix process $\tilde{\boldsymbol{W}}^1$ is φ_1-time changed and the matrix process $\tilde{\boldsymbol{W}}^2$ is φ_2-time changed. Now we construct a new 4×2 matrix process $\bar{\boldsymbol{W}}$ with elements

$$\bar{W}_t^{i,j} = \begin{cases} w^{i,j} + \tilde{W}^{1,i,j}_{\varphi_1(t)} & \text{for } j = 1 \\ w^{i,j} + \tilde{W}^{2,i,j}_{\varphi_2(t)} & \text{for } j = 2, \end{cases} \tag{2.6.26}$$

where $i \in \{1, 2, 3, 4\}$. The trajectory of the eight elements of this 4×2 matrix process are displayed in Fig. 2.6.1. Here we assumed the starting values of the components to be $w^{i,j} = \frac{1}{2}$ for $i \in \{1, 2, 3, 4\}$, $j = 1, 2$ and $\alpha_0^1 = 0.04$, $\alpha_0^2 = 0.05$, $\eta^1 = 0.05$ and $\eta^2 = 0.06$.

The trajectory of a time changed Wishart process is then obtained by the formula $\bar{\boldsymbol{S}}_t = \bar{\boldsymbol{W}}_t^\top \bar{\boldsymbol{W}}_t$, for $t \in [0, \infty)$, see (2.3.51). We illustrate this 2×2 matrix process in Fig. 2.6.2. Note that the diagonal elements of this matrix are

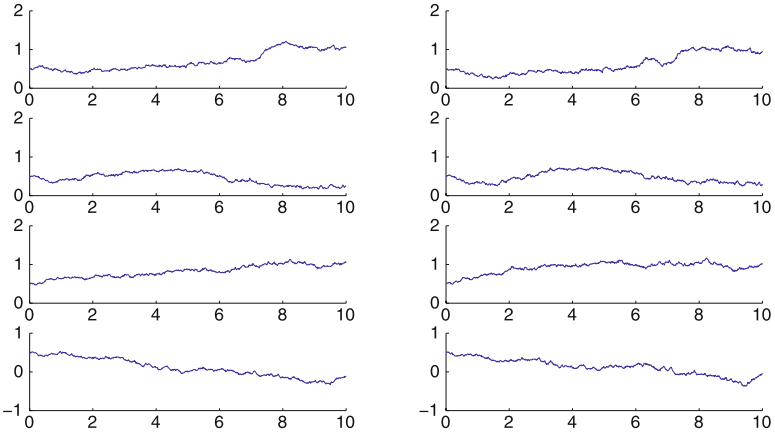

Fig. 2.6.1. Trajectory of the 4×2 matrix time changed Wiener process with independent components in the columns and correlated components in the rows

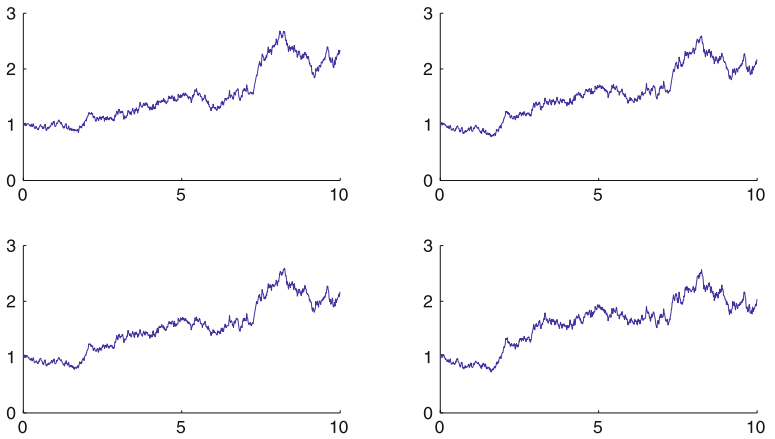

Fig. 2.6.2. Trajectory of the 2×2 time changed Wishart matrix

correlated time changed squared Bessel processes constructed in the following way

$$\bar{S}_t^{1,1} = \left(w^{1,1} + \tilde{W}^{1,1,1}_{\varphi_1(t)}\right)^2 + \left(w^{2,1} + \tilde{W}^{1,2,1}_{\varphi_1(t)}\right)^2 + \left(w^{3,1} + \tilde{W}^{1,3,1}_{\varphi_1(t)}\right)^2 + \left(w^{4,1} + \tilde{W}^{1,4,1}_{\varphi_1(t)}\right)^2,$$

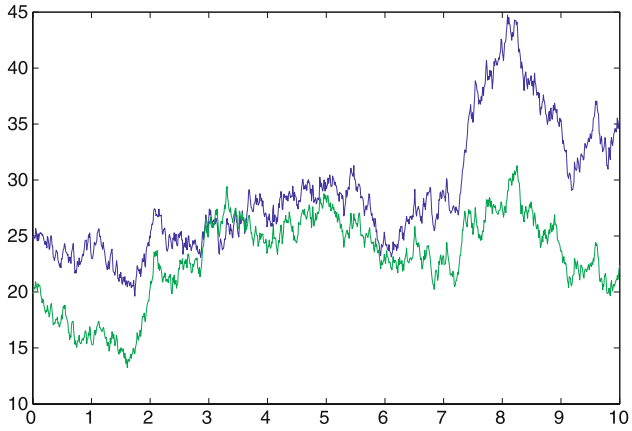

Fig. 2.6.3. Trajectory of two related SR-processes modeling the normalized GOP for two different currencies

$$\begin{aligned}\bar{S}_t^{2,2} = &\left(w^{1,2} + \varrho\,\tilde{W}^{2,1,1}_{\varphi_2(t)} + \sqrt{1-\varrho^2}\,\tilde{W}^{2,1,2}_{\varphi_2(t)}\right)^2 \\ &+ \left(w^{2,2} + \varrho\,\tilde{W}^{2,2,1}_{\varphi_2(t)} + \sqrt{1-\varrho^2}\,\tilde{W}^{2,2,2}_{\varphi_2(t)}\right)^2 \\ &+ \left(w^{3,2} + \varrho\,\tilde{W}^{2,3,1}_{\varphi_2(t)} + \sqrt{1-\varrho^2}\,\tilde{W}^{2,3,2}_{\varphi_2(t)}\right)^2 \\ &+ \left(w^{4,2} + \varrho\,\tilde{W}^{2,4,1}_{\varphi_2(t)} + \sqrt{1-\varrho^2}\,\tilde{W}^{2,4,2}_{\varphi_2(t)}\right)^2, \end{aligned} \tag{2.6.27}$$

where $S_0^{1,1} = \sum_{i=1}^4 (w^{i,1})^2$ and $S_0^{2,2} = \sum_{i=1}^4 (w^{i,2})^2$. These two processes model the discounted GOP in units of the two respective currencies. Let us also illustrate the normalized GOP for these two currencies. For this part of the simulation it is sufficient to work with the above two squared Bessel processes. We construct, according to (2.6.23), two related SR-processes by setting

$$Y_t^1 = \frac{\bar{S}_t^{1,1}}{\alpha_t^1} \quad \text{and} \quad Y_t^2 = \frac{\bar{S}_t^{2,2}}{\alpha_t^2}, \tag{2.6.28}$$

for $t \in [0, \infty)$. These two processes are illustrated with their trajectories in Fig. 2.6.3.

Multi-dimensional Generalized MMM

Let us now generalize the MMM with a time change by introducing a *market activity* process $m = \{m_t, t \in [0, \infty)\}$. That is, we will use time $\tau = \{\tau(t), t \in$

$[0,\infty)\}$, which we call *market time*, to model random variations in market activity. The market time process τ is given by the relation

$$\tau(t) = \tau(0) + \int_0^t m_s ds, \tag{2.6.29}$$

where $\tau_0 > 0$, see Breymann, Kelly & Platen (2006).

In this generalized setting the discounted GOP value $\bar{S}_t^{\delta_*} = \bar{S}^{\delta_*}(\tau(t))$ is modeled in market time τ via the SDE

$$d\bar{S}^{\delta_*}(\tau(t)) = \alpha^{\delta_*}(\tau(t))\, d\tau(t) + \sqrt{\bar{S}^{\delta_*}(\tau(t))\, \alpha^{\delta_*}(\tau(t))}\, dW(\tau(t)), \tag{2.6.30}$$

for $t \in [0,\infty)$, where $W(\cdot)$ is a standard Wiener process in τ-time. The SDE for the normalized GOP $Y(\tau(t)) = \frac{\bar{S}^{\delta_*}(\tau(t))}{\alpha^{\delta_*}(\tau(t))}$ is then

$$dY(\tau(t)) = (1 - \eta Y(\tau(t)))\, d\tau(t) + \sqrt{Y(\tau(t))}\, dW(\tau(t)), \tag{2.6.31}$$

for $t \in [0,\infty)$. This SDE can be rewritten for $Y_t = Y(\tau(t))$ in t-time in the following way

$$dY_t = (1 - \eta Y_t) m_t\, dt + \sqrt{Y_t m_t}\, dW_t, \tag{2.6.32}$$

for $t \in [0,\infty)$, where $W_\cdot$ is a standard Wiener process in t-time.

Now, let the market activity process m be modeled by a linear diffusion process as considered in Sect. 2.4, that is

$$dm_t = \kappa(\bar{m} - m_t)\, dt + \gamma m_t\, d\tilde{W}_t, \tag{2.6.33}$$

for $t \in [0,\infty)$. Here $\tilde{W}_t$ is a standard Wiener process in t-time. Note that the above market activity resembles the squared volatility of an ARCH diffusion asset price, see (1.7.33).

Let us now consider a market with $d+1$ currencies. We can describe the value of the GOP in τ_i-time in the ith currency denomination by the expression

$$S_i^{\delta_*}(\tau_i(t)) = \alpha^i(\tau_i(t)) Y_t^i S_i^i(\tau_i(t)), \tag{2.6.34}$$

where $\alpha^i(\tau_i(t)) = \alpha_0^i \exp\{\eta^i \tau_i(t)\}$ and $S_i^i(\tau_i(t)) = \exp\{r^i \tau_i(t)\}$. Here η^i is the ith net growth rate and r^i the ith short rate as in the case of the stylized MMM. The ith normalized GOP satisfies the SDE

$$dY_t^i = (1 - \eta^i Y_t^i) m_t^i\, dt + \sqrt{Y_t^i m_t^i}\, dW_t^i \tag{2.6.35}$$

for $t \in [0,\infty)$, with $Y^i(\tau_i(0)) > 0$, where $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, W_t^2, \ldots, W_t^{d+1})^\top, t \in [0,\infty)\}$ is a vector of correlated Wiener processes. Additionally, the market activity process $\boldsymbol{m} = \{\boldsymbol{m}_t = (m_t^1, m_t^2, \ldots, m_t^{d+1})^\top, t \in [0,\infty)\}$ is a $d+1$-dimensional linear diffusion process as discussed in Sect. 2.4.

Let us now illustrate the simulation of a two currency market under the generalized MMM. Here we use the market activity given by the correlated

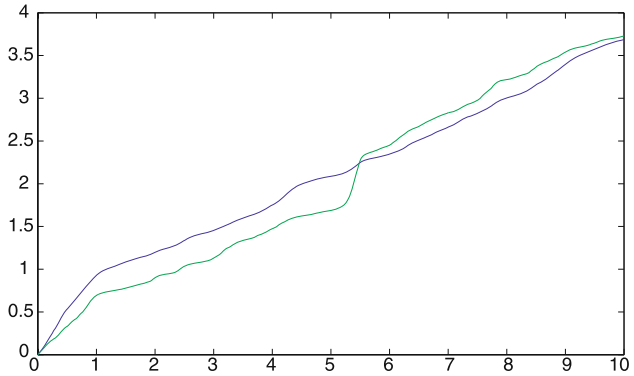

Fig. 2.6.4. Two related market times τ_1 and τ_2 obtained from the linear diffusions in Fig. 2.4.2

linear diffusion processes in Fig. 2.4.2. By approximation of the integral in (2.6.29) we obtain two correlated market times which we display in Fig. 2.6.4. Given these two nondecreasing processes we then proceed with the simulation as in the case of the stylized MMM. However, we use now the time change

$$\bar{\varphi}_i(t) = \varphi_i(\tau_i(t)), \tag{2.6.36}$$

for $i \in \{1, 2\}$ to include the effect of random market activity. This leads to two correlated squared Bessel processes in two correlated times analogous to (2.6.26)–(2.6.27) given by

$$\begin{aligned} \bar{S}_t^{1,1} = &\left(w^{1,1} + \tilde{W}^{1,1,1}_{\bar{\varphi}_1(t)}\right)^2 + \left(w^{2,1} + \tilde{W}^{1,2,1}_{\bar{\varphi}_1(t)}\right)^2 \\ &+ \left(w^{3,1} + \tilde{W}^{1,3,1}_{\bar{\varphi}_1(t)}\right)^2 + \left(w^{4,1} + \tilde{W}^{1,4,1}_{\bar{\varphi}_1(t)}\right)^2, \end{aligned} \tag{2.6.37}$$

$$\begin{aligned} \bar{S}_t^{2,2} = &\left(w^{1,2} + \varrho\, \tilde{W}^{2,1,1}_{\bar{\varphi}_2(t)} + \sqrt{1-\varrho^2}\, \tilde{W}^{2,1,2}_{\bar{\varphi}_2(t)}\right)^2 \\ &+ \left(w^{2,2} + \varrho\, \tilde{W}^{2,2,1}_{\bar{\varphi}_2(t)} + \sqrt{1-\varrho^2}\, \tilde{W}^{2,2,2}_{\bar{\varphi}_2(t)}\right)^2 \\ &+ \left(w^{3,2} + \varrho\, \tilde{W}^{2,3,1}_{\bar{\varphi}_2(t)} + \sqrt{1-\varrho^2}\, \tilde{W}^{2,3,2}_{\bar{\varphi}_2(t)}\right)^2 \\ &+ \left(w^{4,2} + \varrho\, \tilde{W}^{2,4,1}_{\bar{\varphi}_2(t)} + \sqrt{1-\varrho^2}\, \tilde{W}^{2,4,2}_{\bar{\varphi}_2(t)}\right)^2, \end{aligned} \tag{2.6.38}$$

where $S_0^{1,1} = \sum_{i=1}^4 (w^{i,1})^2$ and $S_0^{2,2} = \sum_{i=1}^4 (w^{i,2})^2$. We display the entire Wishart matrix process in τ-time in Fig. 2.6.5, whose diagonal elements are as

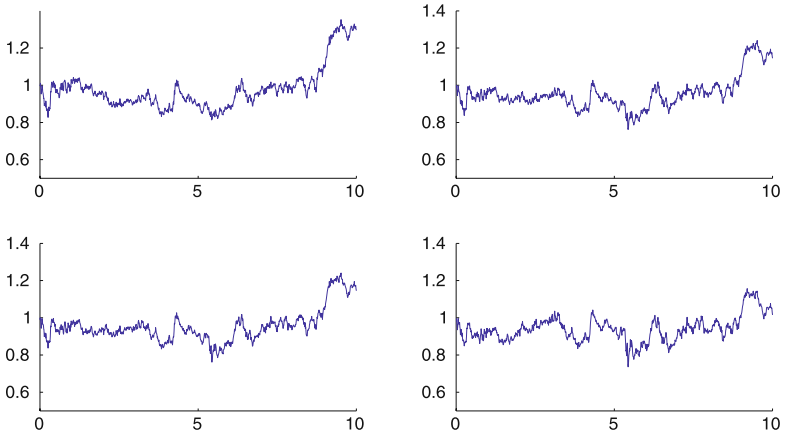

Fig. 2.6.5. Trajectory of the 2×2 time changed Wishart process

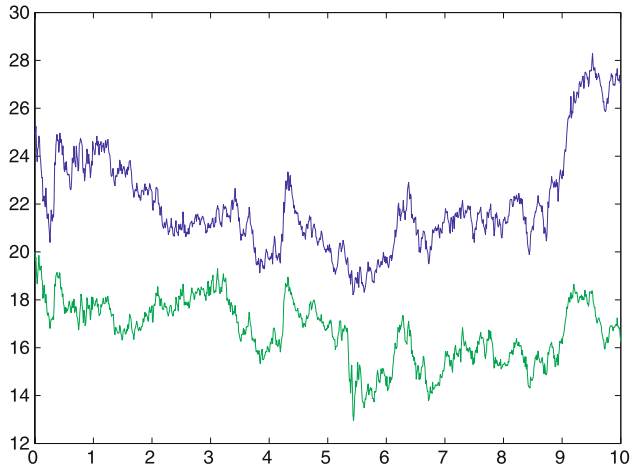

Fig. 2.6.6. Trajectory of two correlated SR-processes modeling the normalized GOP for two different currencies under the generalized MMM

in (2.6.37) and (2.6.38). Additionally, we construct two related SR-processes as

$$Y_t^1 = \frac{\bar{S}_t^{1,1}}{\bar{\alpha}^1(\tau_1(t))} \quad \text{and} \quad Y_t^2 = \frac{\bar{S}_t^{2,2}}{\bar{\alpha}^2(\tau_2(t))}, \tag{2.6.39}$$

for $t \in [0, \infty)$. These two processes are illustrated in Fig. 2.6.6. Moreover, in Fig. 2.6.7 we display the quadratic variations when multiplied by 4 of the

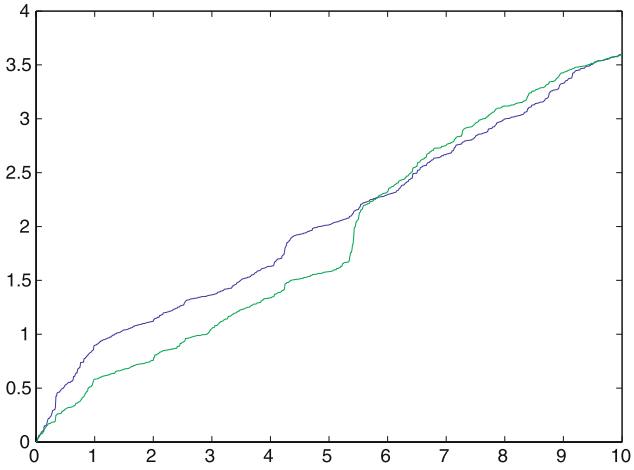

Fig. 2.6.7. Quadratic variation×4 of the square root of the SR-process from Fig. 2.6.6

square root of the SR-processes from Fig. 2.6.6. These should approximately be equal to the corresponding τ-times already displayed in Fig. 2.6.4, since

$$[\sqrt{Y^i}]_t = \frac{1}{4}\tau_i(t) \tag{2.6.40}$$

for $i \in \{1, 2\}$. The visualization confirms the usefulness of the presented almost exact simulation method for the generalized MMM.

The above presented exact and almost exact simulation methods for multi-dimensional diffusions lead to highly accurate scenario simulations. Since no recursive scheme is involved the minor errors are not propagated and the methods are also reliable over long periods of time. The numerical stability problems that we will highlight in Chap. 14 for recursive discrete-time approximations of SDEs are avoided. We remark that the effect of jumps on the above considered type of dynamics can be introduced via a jump-adapted time discretization, see Platen (1982a) and will be discussed in Sect. 8.6.

2.7 Functionals of Solutions of SDEs

Explicit formulas for functionals of solutions of SDEs are of particular interest. As we will see in Chap. 3, pricing rules are usually expressed via pricing formulas that have the form of conditional expectations. For Markovian state variables these conditional expectations lead to pricing functions that satisfy certain *partial differential equations* (PDEs), or more generally, *partial*

integro differential equations (PIDEs) when jumps are involved. The link between the conditional expectations and respective PDEs can be interpreted as an application of the, so-called, *Feynman-Kac formula*. Below we formulate the Feynman-Kac formula in various ways. For a wide range of models this formula provides the PDEs and PIDES that characterize corresponding pricing functions of derivatives. Furthermore, this section presents some results on transition probability densities, changes of measures, the Bayes rule, the Girsanov transformation and the Black-Scholes option pricing formula. These are useful when searching for analytic formulas of derivative prices and other functionals.

SDE for Some Factor Process

We consider a fixed time horizon $T \in (0, \infty)$ and a d-dimensional Markovian factor process $\boldsymbol{X}^{t,\boldsymbol{x}} = \{\boldsymbol{X}_s^{t,\boldsymbol{x}},\, s \in [t,T]\}$, which satisfies the vector SDE

$$d\boldsymbol{X}_s^{t,\boldsymbol{x}} = \boldsymbol{a}(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds + \sum_{k=1}^{m} \boldsymbol{b}^{\boldsymbol{k}}(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, dW_s^k \tag{2.7.1}$$

for $s \in [t,T]$ with initial value $\boldsymbol{X}_t^{t,x} = \boldsymbol{x} \in \Re^d$ at time $t \in [0,T]$. The process $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, \ldots, W_t^m)^\top,\ t \in [0,T]\}$ is an m-dimensional standard Wiener process on a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$. The process $\boldsymbol{X}^{t,\boldsymbol{x}}$ has a drift coefficient $\boldsymbol{a}(\cdot,\cdot)$ and diffusion coefficients $\boldsymbol{b}^{\boldsymbol{k}}(\cdot,\cdot)$, $k \in \{1,2,\ldots,m\}$. In general, $\boldsymbol{a} = (a^1, \ldots, a^d)^\top$ and $\boldsymbol{b}^{\boldsymbol{k}} = (b^{1,k}, \ldots, b^{d,k})^\top$, $k \in \{1,2,\ldots,m\}$, represent vector valued functions on $[0,T] \times \Re^d$ into $\Re^d$, and we assume that a pathwise unique solution of the SDE (2.7.1) exists. The components of the SDE (2.7.1) can be the factors in a financial market model. In the following it will not matter which pricing approach one chooses. The task will be to evaluate simply some conditional expectations.

Terminal Payoff

Let us discuss the case of a European option, where we have a terminal payoff $H(\boldsymbol{X}_T^{t,\boldsymbol{x}})$ at the maturity date T with some given payoff function $H : \Re^d \to [0,\infty)$ such that

$$E(|H(\boldsymbol{X}_T^{t,\boldsymbol{x}})|) < \infty. \tag{2.7.2}$$

We can then introduce the pricing function $u : [0,T] \times \Re^d \to [0,\infty)$ as the conditional expectation

$$u(t,\boldsymbol{x}) = E\left(H(\boldsymbol{X}_T^{t,\boldsymbol{x}}) \,\middle|\, \mathcal{A}_t\right) \tag{2.7.3}$$

for $(t,\boldsymbol{x}) \in [0,T] \times \Re^d$. The *Feynman-Kac formula* for this payoff refers to the fact that under sufficient regularity on $\boldsymbol{a}, \boldsymbol{b^1}, \ldots, \boldsymbol{b^m}$ and H the function $u : (0,T) \times \Re^d \to [0,\infty)$ satisfies the PDE

$$L^0 u(t,\boldsymbol{x}) = \frac{\partial u(t,\boldsymbol{x})}{\partial t} + \sum_{i=1}^{d} a^i(t,\boldsymbol{x}) \frac{\partial u(t,\boldsymbol{x})}{\partial x^i}$$
$$+ \frac{1}{2} \sum_{i,k=1}^{d} \sum_{j=1}^{m} b^{i,j}(t,\boldsymbol{x})\, b^{k,j}(t,\boldsymbol{x}) \frac{\partial^2 u(t,\boldsymbol{x})}{\partial x^i \partial x^k}$$
$$= 0 \tag{2.7.4}$$

for $(t,\boldsymbol{x}) \in (0,T) \times \Re^d$ with terminal condition

$$u(T,\boldsymbol{x}) = H(\boldsymbol{x}) \tag{2.7.5}$$

for $\boldsymbol{x} \in \Re^d$. Equation (2.7.4) is also called the *Kolmogorov backward equation*.

Note that in general, one needs also to specify the behavior of the solution of the PDE at its boundaries. In many cases this is not adding any extra information as is, for instance, the case under the Merton model in Sect. 3.5. However, in some cases, for instance under the minimal market model, as will be discussed in Sect. 3.6, it can be crucial. In financial applications it is important to set the boundaries such that strong arbitrage, in the sense as will be described in Definition 3.3.2, is excluded. This means that a derivative price needs to have an absorbing boundary condition at zero, where when it reaches zero, it keeps its value afterwards at zero.

The above type of European payoff will be covered by a version of the Feynman-Kac formula that we will present below. Under real world pricing the above version of the Feynman-Kac formula allows the calculation of the benchmarked pricing function under the real world probability measure. For instance, it can also be applied to determine the discounted pricing function under risk neutral pricing when the expectation can be taken of the discounted payoff under an equivalent risk neutral probability measure.

Discounted Payoff

We now generalize the above payoff function by discounting it, using a given *discount rate process* r, which is obtained as a function of the given vector diffusion process $\boldsymbol{X}^{t,\boldsymbol{x}}$, that is $r : [0,T] \times \Re^d \to \Re$. For instance, in a risk neutral setting the discount rate is given by the short term interest rate.

Over the period $[t,T]$ we consider for the *discounted payoff*

$$\exp\left\{ - \int_t^T r(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds \right\} H(\boldsymbol{X}_T^{t,\boldsymbol{x}})$$

the pricing function

$$u(t,\boldsymbol{x}) = E\left(\exp\left\{ - \int_t^T r(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds \right\} H(\boldsymbol{X}_T^{t,\boldsymbol{x}}) \,\middle|\, \mathcal{A}_t \right) \tag{2.7.6}$$

for $(t, \boldsymbol{x}) \in [0, T] \times \Re^d$. As will be shown below, it follows rather generally that the pricing function u satisfies the PDE

$$L^0 u(t, \boldsymbol{x}) = r(t, \boldsymbol{x})\, u(t, \boldsymbol{x}) \tag{2.7.7}$$

for $(t, \boldsymbol{x}) \in (0, T) \times \Re^d$ with terminal condition

$$u(T, \boldsymbol{x}) = H(\boldsymbol{x}) \tag{2.7.8}$$

for $\boldsymbol{x} \in \Re^d$. Here the PDE operator L^0 is given as in (2.7.4).

Terminal Payoff and Payoff Rate

Now, we add to the above discounted payoff structure some payoff stream that continuously pays with a *payoff rate* $g : [0, T] \times \Re^d \to [0, \infty)$ some amount per unit of time. This allows to model, for instance, continuous dividend payments for a share or continuous interest payments in a savings account. Also fees and insurance premia can be captured in this way. The corresponding *discounted payoff with payoff rate* is then of the form

$$\exp\left\{-\int_t^T r(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds\right\} H(\boldsymbol{X}_T^{t,\boldsymbol{x}}) + \int_t^T \exp\left\{-\int_t^s r(z, \boldsymbol{X}_z^{t,\boldsymbol{x}})\, dz\right\} g(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds,$$

which leads to the pricing function

$$\begin{aligned} u(t, \boldsymbol{x}) = E\Bigg(&\exp\left\{-\int_t^T r(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds\right\} H(\boldsymbol{X}_T^{t,\boldsymbol{x}}) \\ &+ \int_t^T \exp\left\{-\int_t^s r(z, \boldsymbol{X}_z^{t,\boldsymbol{x}})\, dz\right\} g(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds \,\Big|\, \mathcal{A}_t\Bigg) \end{aligned} \tag{2.7.9}$$

for $(t, \boldsymbol{x}) \in [0, T] \times \Re^d$. As will follow below, this pricing function satisfies the PDE

$$L^0 u(t, \boldsymbol{x}) + g(t, \boldsymbol{x}) = r(t, \boldsymbol{x})\, u(t, \boldsymbol{x}) \tag{2.7.10}$$

for $(t, \boldsymbol{x}) \in (0, T) \times \Re^d$ with terminal condition

$$u(T, \boldsymbol{x}) = H(\boldsymbol{x}) \tag{2.7.11}$$

for $\boldsymbol{x} \in \Re^d$. As mentioned earlier, for certain dynamics boundary conditions may have to be added for completeness.

SDE with Jumps

We consider now an SDE with jumps describing the dynamics of the underlying factor process. Let Γ denote an open, connected subset of $\Re^d$ and

$T \in (0,\infty)$ a fixed time horizon. We consider for a d-dimensional process $\boldsymbol{X}^{t,\boldsymbol{x}} = \{\boldsymbol{X}_s^{t,\boldsymbol{x}},\, s \in [t,T]\}$, see (1.5.19), the vector SDE

$$d\boldsymbol{X}_s^{t,\boldsymbol{x}} = \boldsymbol{a}(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds + \sum_{k=1}^{m} \boldsymbol{b}^{\boldsymbol{k}}(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, dW_s^k$$

$$+ \sum_{j=1}^{\ell} \int_{\mathcal{E}} \boldsymbol{c}^{\boldsymbol{j}}(v, s-, \boldsymbol{X}_{s-}^{t,\boldsymbol{x}})\, p_{\varphi_j}^j(dv, ds) \tag{2.7.12}$$

for $t \in [0,T]$ and $s \in [t,T]$ with value

$$\boldsymbol{X}_t^{t,\boldsymbol{x}} = \boldsymbol{x} \tag{2.7.13}$$

at time t and for $\boldsymbol{x} \in \Gamma$, see (1.6.16). Here $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, \ldots, W_t^m)^\top,\, t \in [0,T]\}$ is again an m-dimensional standard Wiener process on a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$. Furthermore, $p_{\varphi_j}^j(\cdot,\cdot)$ denotes a Poisson measure, $j \in \{1,2,\ldots,\ell\}$, as introduced in Sect. 1.1, satisfying condition (1.1.33). Here $\boldsymbol{a} = (a^1, \ldots, a^d)^\top$ and $\boldsymbol{b}^{\boldsymbol{k}} = (b^{1,k}, \ldots, b^{d,k})^\top$, $k \in \{1,2,\ldots,m\}$, are vector valued functions from $[0,T] \times \Gamma$ into $\Re^d$ and $\boldsymbol{c}^{\boldsymbol{j}} = (c^{1,j}, \ldots, c^{d,j})^\top$, $j \in \{1,2,\ldots,\ell\}$, is a vector valued function on $\mathcal{E} \times [0,T] \times \Gamma$, where $\mathcal{E} = \Re \backslash \{0\}$.

For the above payoff structure, with discounted terminal payoff and a given payoff rate, the pricing function is of the form (2.7.9). As will be detailed below, under appropriate conditions u satisfies the PIDE

$$L^0\, u(t,\boldsymbol{x}) + g(t,\boldsymbol{x}) = r(t,\boldsymbol{x})\, u(t,\boldsymbol{x}) \tag{2.7.14}$$

for $(t,\boldsymbol{x}) \in (0,T)$ with terminal condition

$$u(T,\boldsymbol{x}) = H(\boldsymbol{x}) \tag{2.7.15}$$

for $\boldsymbol{x} \in \Re^d$. The corresponding operator L^0 is more general than the one in (2.7.4) and given in the form

$$L^0 u(t,\boldsymbol{x}) = \sum_{i=1}^{d} a^i(t,\boldsymbol{x})\, \frac{\partial u(t,\boldsymbol{x})}{\partial x^i} + \frac{1}{2} \sum_{i,k=1}^{d} \sum_{j=1}^{m} b^{i,j}(t,\boldsymbol{x})\, b^{k,j}(t,\boldsymbol{x})\, \frac{\partial^2 u(t,\boldsymbol{x})}{\partial x^i\, \partial x^k}$$

$$+ \frac{\partial u(t,\boldsymbol{x})}{\partial t} + \sum_{j=1}^{\ell} \int_{\mathcal{E}} \big[u(s, x^1 + c^{1,j}(v,s,\boldsymbol{x}), \ldots, x^d + c^{d,j}(v,s,\boldsymbol{x}))$$

$$- u(s, x^1, \ldots, x^d)\big]\, \varphi_j(dv). \tag{2.7.16}$$

Here we abuse slightly our notation by writing for $u(s,(x^1,\ldots,x^d)^\top)$ also $u(s,x^1,\ldots,x^d)$. Note that due to the jumps an extra integral term appears in (2.7.16) as a consequence of the Itô formula with jumps, see (1.5.15) and (1.5.20). Of course, also here some boundary conditions may have to be added to complete the characterization of the function u via the PIDE.

Payoff with First Exit Time

Derivatives like barrier options have a, so-called, *continuation region* Φ, which is an open connected subset of $[0,T] \times \Gamma$. The holder of an option continues to receive payments as long as the process $\boldsymbol{X}^{t,\boldsymbol{x}}$ stays in the continuation region Φ. For instance, in the case of a, so-called, *knock-out-barrier option* this would mean that $\boldsymbol{X}_s^{t,\boldsymbol{x}}$ has to stay below a given critical barrier to receive the terminal payment. To make this precise, we define the *first exit time* τ_Φ^t from Φ after t as

$$\tau_\Phi^t = \inf\{s \in [t,T] : (s, \boldsymbol{X}_s^{t,\boldsymbol{x}}) \notin \Phi\}, \tag{2.7.17}$$

which is a stopping time, see (1.2.11).

Consider now a general payoff structure with *terminal payoff function* $H : (0,T] \times \Gamma \to [0,\infty)$ for payments at time τ_Φ^t, a *payoff rate* $g : [0,T] \times \Gamma \to [0,\infty)$ for incremental payments during the time period $[t, \tau_\Phi^t)$ and a *discount rate* $r : [0,T] \times \Gamma \to \Re$. Assume that the process $\boldsymbol{X}^{t,\boldsymbol{x}}$ does not explode or leave Γ before the terminal time T. We then define the *pricing function* $u : \Phi \to [0,\infty)$ by

$$u(t,\boldsymbol{x}) = E\left(H(\tau_\Phi^t, \boldsymbol{X}_{\tau_\Phi^t}^{t,\boldsymbol{x}}) \exp\left\{ -\int_t^{\tau_\Phi^t} r(s, \boldsymbol{X}_s^{t,\boldsymbol{x}})\, ds \right\} \right.$$
$$\left. + \int_t^{\tau_\Phi^t} g(s, \boldsymbol{X}_s^{t,\boldsymbol{x}}) \exp\left\{ -\int_t^s r(z, \boldsymbol{X}_u^{t,\boldsymbol{x}})\, dz \right\} ds \,\middle|\, \mathcal{A}_t \right) \tag{2.7.18}$$

for $(t,\boldsymbol{x}) \in \Phi$.

For the formulation of the resulting PIDE of the function u we use the operator L^0 given in (2.7.16). Under sufficient regularity of Φ, $\boldsymbol{a}$, $\boldsymbol{b^1}$, ..., $\boldsymbol{b^m}$, $\boldsymbol{c^1}, \ldots, \boldsymbol{c^\ell}$, H, g, φ_1, ..., φ_ℓ and r one can show by application of the Itô formula (1.5.15) and some martingale argument that the pricing function u satisfies the PIDE

$$L^0 u(t,\boldsymbol{x}) + g(t,\boldsymbol{x}) = r(t,\boldsymbol{x})\, u(t,\boldsymbol{x}) \tag{2.7.19}$$

for $(t,\boldsymbol{x}) \in \Phi$ with boundary condition

$$u(t,\boldsymbol{x}) = H(t,\boldsymbol{x}) \tag{2.7.20}$$

for $(t,\boldsymbol{x}) \in ((0,T] \times \Gamma) \backslash \Phi$. This result links the functional (2.7.18) to the PIDE (2.7.19)–(2.7.20) and is often called a *Feynman-Kac formula*.

Generalized Feynman-Kac Formula

For a rather general situation, where $\Phi = (0,T) \times \Gamma$ and assuming no jumps, that is $\boldsymbol{c^1} = \ldots = \boldsymbol{c^\ell} = 0$ and $\tau_\Phi^t = T$, let us now formulate sufficient conditions that ensure that the Feynman-Kac formula holds, see Heath & Schweizer (2000) and Platen & Heath (2006).

(A) The drift coefficient $\boldsymbol{a}$ and diffusion coefficients $\boldsymbol{b^k}$, $k \in \{1, 2, \ldots, m\}$, are assumed to be on $[0, T] \times \Gamma$ locally Lipschitz-continuous in $\boldsymbol{x}$, uniformly in t. That is, for each compact subset Γ^1 of Γ there exists a constant $K_{\Gamma^1} < \infty$ such that

$$|\boldsymbol{a}(t, \boldsymbol{x}) - \boldsymbol{a}(t, \boldsymbol{y})| + \sum_{k=1}^{m} |\boldsymbol{b^k}(t, \boldsymbol{x}) - \boldsymbol{b^k}(t, \boldsymbol{y})| \le K_{\Gamma^1} |\boldsymbol{x} - \boldsymbol{y}| \tag{2.7.21}$$

for all $t \in [0, T]$ and $\boldsymbol{x}, \boldsymbol{y} \in \Gamma^1$.

(B) For all $(t, \boldsymbol{x}) \in [0, T) \times \Gamma$ the solution $\boldsymbol{X}^{t,\boldsymbol{x}}$ of (2.7.12) neither explodes nor leaves Γ before T, that is

$$P\left(\sup_{t \le s \le T} |\boldsymbol{X}_s^{t,\boldsymbol{x}}| < \infty \right) = 1 \tag{2.7.22}$$

and

$$P(\boldsymbol{X}_s^{t,\boldsymbol{x}} \in \Gamma \text{ for all } s \in [t, T]) = 1. \tag{2.7.23}$$

(C) There exists an increasing sequence $(\Gamma_n)_{n \in \mathcal{N}}$ of bounded, open and connected domains of Γ such that $\cup_{n=1}^{\infty} \Gamma_n = \Gamma$, and for each $n \in \mathcal{N}$ the PDE

$$L^0 u_n(t, \boldsymbol{x}) + g(t, \boldsymbol{x}) = r(t, \boldsymbol{x})\, u_n(t, \boldsymbol{x}) \tag{2.7.24}$$

has a unique solution u_n in the sense of Friedman (1975) on $(0, T) \times \Gamma_n$ with boundary condition

$$u_n(t, \boldsymbol{x}) = u(t, \boldsymbol{x}) \tag{2.7.25}$$

on $((0, T) \times \partial\Gamma_n) \cup (\{T\} \times \Gamma_n)$, where $\partial\Gamma_n$ denotes the boundary of Γ_n.

(D) The process $b^{i,k}(\cdot, \boldsymbol{X}_\cdot) \frac{\partial u(\cdot, \boldsymbol{X}_\cdot)}{\partial x^i}$ is measurable and square integrable on $[0, T]$ for all $i \in \{1, 2, \ldots, d\}$ and $k \in \{1, 2, \ldots, m\}$.

The proof of the following theorem is given in Platen & Heath (2006).

Theorem 2.7.1. *In the case without jumps and under the conditions (A), (B), (C) and (D), the function u given by (2.7.18) is the unique solution of the PDE (2.7.19) with boundary condition (2.7.20), where u is differentiable with respect to t and twice differentiable with respect to the components of $\boldsymbol{x}$.*

Condition (A) is satisfied if, for instance, $\boldsymbol{a}$ and $\boldsymbol{b} = (\boldsymbol{b^1}, \ldots, \boldsymbol{b^m})$ are differentiable in $\boldsymbol{x}$ on the open set $(0, T) \times \Gamma$ with derivatives that are continuous on $[0, T] \times \Gamma$.

To establish condition (B) one needs to exploit specific properties of the process $\boldsymbol{X}^{t,\boldsymbol{x}}$ given by the SDE (2.7.12).

Condition (C) can be shown to be implied by the following assumptions:

(C1) There exists an increasing sequence $(\Gamma_n)_{n \in \mathcal{N}}$ of bounded, open and connected subdomains of Γ with $\Gamma_n \cup \partial\Gamma_n \subset \Gamma$ such that $\cup_{n=1}^{\infty} \Gamma_n = \Gamma$, and each Γ_n has a twice differentiable boundary $\partial\Gamma_n$.

(C2) For each $n \in \mathcal{N}$ the functions $\boldsymbol{a}$ and $\boldsymbol{b}\boldsymbol{b}^\top$ are uniformly Lipschitz-continuous on $[0,T] \times (\Gamma_n \cup \partial\Gamma_n)$.

(C3) For each $n \in \mathcal{N}$ the function $\boldsymbol{b}(t,\boldsymbol{x})\boldsymbol{b}(t,\boldsymbol{x})^\top$ is uniformly elliptic on $\Re^d$ for $(t,\boldsymbol{x}) \in [0,T] \times \Gamma_n$, that is, there exists a $\delta_n > 0$ such that

$$\boldsymbol{y}^\top \boldsymbol{b}(t,\boldsymbol{x})\, \boldsymbol{b}(t,\boldsymbol{x})^\top \boldsymbol{y} \geq \delta_n \, |\boldsymbol{y}|^2 \tag{2.7.26}$$

for all $\boldsymbol{y} \in \Re^d$.

(C4) For each $n \in \mathcal{N}$ the functions r and g are uniformly *Hölder-continuous* on $[0,T] \times (\Gamma_n \cup \partial\Gamma_n)$, that is, there exists a constant $\bar{K}_n$ and an exponent $q_n > 0$ such that

$$|r(t,\boldsymbol{x}) - r(t,\boldsymbol{y})| + |g(t,\boldsymbol{x}) - g(t,\boldsymbol{y})| \leq \bar{K}_n \, |\boldsymbol{x} - \boldsymbol{y}|^{q_n} \tag{2.7.27}$$

for $t \in [0,T]$ and $\boldsymbol{x}, \boldsymbol{y} \in (\Gamma_n \cup \partial\Gamma_n)$.

(C5) For each $n \in \mathcal{N}$ the function u is finite and continuous on $([0,T] \times \partial\Gamma_n) \cup (\{T\} \times (\Gamma_n \cup \partial\Gamma_n))$.

Condition (D) is satisfied, for instance, when

$$\int_0^T E\left(\left(b^{i,k}(t,\boldsymbol{X}_t)\,\frac{\partial u(t,\boldsymbol{X}_t)}{\partial x^i}\right)^2\right) dt < \infty$$

for all $i \in \{1,2,\ldots,d\}$ and $k \in \{1,2,\ldots,m\}$. This condition ensures that the process $u(\cdot,\boldsymbol{X}_\cdot)$ is a martingale and that the PDE (2.7.19)–(2.7.20) has a unique solution.

Note that in the case when local Lipschitz continuity is not guaranteed, one may have to specify particular boundary conditions to obtain an appropriate description of the pricing function. This is a consequence of the fact that strict local martingales may drive the factor dynamics. These need extra care when defining the behavior of PDE solutions at boundaries.

Kolmogorov Equations

When the drift coefficient $a(\cdot)$ and diffusion coefficient $b(\cdot)$ of the solution of a scalar SDE without jumps are appropriate functions, then the corresponding transition probability density $p(s,x;t,y)$ satisfies a certain PDE. This is the *Kolmogorov forward equation* or *Fokker-Planck equation*

$$\frac{\partial p(s,x;t,y)}{\partial t} + \frac{\partial}{\partial y}\left\{a(t,y)\, p(s,x;t,y)\right\} - \frac{1}{2}\,\frac{\partial^2}{\partial y^2}\left\{b^2(t,y)\, p(s,x;t,y)\right\} = 0, \tag{2.7.28}$$

for (s,x) fixed. However, $p(s,x;t,y)$ satisfies also the *Kolmogorov backward equation*

$$\frac{\partial p(s,x;t,y)}{\partial s} + a(s,x)\frac{\partial p(s,x;t,y)}{\partial x} + \frac{1}{2}\, b^2(s,x)\frac{\partial^2 p(s,x;t,y)}{\partial x^2} = 0, \tag{2.7.29}$$

for (t, y) fixed. Obviously, the *initial condition* for both PDEs is given by the *Dirac delta function*

$$p(s, x; s, y) = \delta(y - x) = \begin{cases} \infty & \text{for} \quad y = x \\ 0 & \text{for} \quad y \neq x, \end{cases} \tag{2.7.30}$$

where

$$\int_{-\infty}^{\infty} \delta(y - x)\, dy = 1 \tag{2.7.31}$$

for given x. In the case of an SDE with jumps (2.7.28) and (2.7.29) become corresponding PIDEs. Of course, the Kolmogorov equations have their obvious multi-dimensional counterparts. In Sect. 2.2 we have given various examples of explicit transition probability densities. The PDEs and PIDEs that we obtain via the Feynman-Kac formula are often called Kolmogorov backward equations.

Stationary Densities

Let us consider solutions of SDEs which provide models for financial quantities that can evolve into some equilibrium. Obviously, we restrict then the class of processes that we consider. For example, such equilibria can be modeled by the Ornstein-Uhlenbeck process or the square root process, see Chap. 1. The corresponding transition probability densities converge over long periods of time towards corresponding stationary densities, see Sect. 2.2.

More precisely, for a diffusion process that permits some equilibrium its *stationary density* $\bar{p}(y)$ is defined as the solution of the integral equation

$$\bar{p}(y) = \int_{-\infty}^{\infty} p(s, x; t, y)\, \bar{p}(x)\, dx$$

for $t \in [0, \infty)$, $s \in [0, t]$ and $y \in \Re$. This means, if one starts with the stationary density, then one obtains again the stationary density as the probability density of the process after any given time period. A stationary solution of an SDE is, therefore, obtained when the corresponding process starts with its stationary density.

It is important to know when the solution of an SDE has a stationary density. One can identify the stationary density $\bar{p}$ by noting that it satisfies the corresponding stationary, or time-independent, Kolmogorov forward equation, see (2.7.28). In the case without jumps this *stationary Kolmogorov forward equation* reduces to the ordinary differential equation (ODE)

$$\frac{d}{dy}\left(a(y)\, \bar{p}(y)\right) - \frac{1}{2}\, \frac{d^2}{dy^2}\left(b^2(y)\, \bar{p}(y)\right) = 0 \tag{2.7.32}$$

with drift $a(x) = a(s, x)$ and diffusion coefficient $b(x) = b(s, x)$. Consequently, it is necessary that equation (2.7.32) is satisfied to ensure that a diffusion has

a stationary density. There is an obvious multi-dimensional PDE generalizing (2.7.32). In the case with jumps a corresponding integral term needs to be added in (2.7.32). We assume in the following that a unique stationary density exists for the process considered.

Note that since $\bar{p}$ is a probability density it must satisfy the relation

$$\int_{-\infty}^{\infty} \bar{p}(y)\, dy = 1. \tag{2.7.33}$$

One can identify for a large class of scalar stationary diffusion processes the analytic form of their stationary density $\bar{p}(y)$. It is straightforward to check that the explicit expression

$$\bar{p}(y) = \frac{C}{b^2(y)} \exp\left\{2 \int_{y_0}^{y} \frac{a(u)}{b^2(u)}\, du\right\} \tag{2.7.34}$$

satisfies the equation (2.7.32) for $y \in \Re$ with some fixed value $y_0 \in \Re$. Here y_0 is some appropriately chosen point so that (2.7.34) makes sense. The constant C can be obtained from the normalization condition (2.7.33).

Ergodicity of a Process

Ergodicity is a concept that is somehow related to stationarity. It does not require the process to start with an initial random variable that already has the stationary density of the process. Ergodicity can be conveniently defined for processes with stationary densities. A process $X = \{X_t,\ t \in [0, \infty)\}$ is called *ergodic* if it has a stationary density $\bar{p}$ and

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(X_t)\, dt = \int_{-\infty}^{\infty} f(x)\, \bar{p}(x)\, dx, \tag{2.7.35}$$

for all bounded measurable functions $f : \Re \to \Re$. That is, the limit as $T \to \infty$ of the random time average, specified on the left hand side of relation (2.7.35), equals the spatial average with respect to $\bar{p}$, as given on the right hand side of (2.7.35).

For a scalar SDE with drift $a(\cdot)$ and diffusion coefficient $b(\cdot)$ let us introduce the *scale measure* $s : \Re \to \Re^+$ given by

$$s(x) = \exp\left\{-2 \int_{y_0}^{x} \frac{a(y)}{b^2(y)}\, dy\right\} \tag{2.7.36}$$

for $x \in \Re$ with y_0 as in (2.7.34). For instance, in Borodin & Salminen (2002) one can find the following result:

A diffusion process with scale measure $s(\cdot)$ satisfying the following two properties:

$$\int_{y_0}^{\infty} s(x)\, dx = \int_{-\infty}^{y_0} s(x)\, dx = \infty \tag{2.7.37}$$

and

$$\int_{-\infty}^{\infty} \frac{1}{s(x)\, b^2(x)}\, dx < \infty \tag{2.7.38}$$

is ergodic and its stationary density $\bar{p}$ is given by the expression (2.7.34). Of course, one can define the notion of ergodicity also for multi-dimensional processes and in the presence of jumps.

Change of Probability Measure

The ability to change measures before applying the Feynman-Kac formula can be powerful, not only in derivative pricing but also for many other tasks, including estimation, filtering and variance reduction, as we will see later in the book.

We denote by $\boldsymbol{W} = \{\boldsymbol{W}_t = (W_t^1, \ldots, W_t^m)^\top,\ t \in [0,T]\}$ an m-dimensional standard Wiener process on a filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ with $\mathcal{A}_0$ being the trivial σ-algebra, augmented by the sets of zero probability. For an $\underline{\mathcal{A}}$-predictable m-dimensional stochastic process $\boldsymbol{\theta} = \{\boldsymbol{\theta}_t = (\theta_t^1, \ldots, \theta_t^m)^\top,\ t \in [0,T]\}$ with

$$\int_0^T \sum_{i=1}^m (\theta_t^i)^2\, dt < \infty \tag{2.7.39}$$

almost surely, let us assume that the strictly positive *Radon-Nikodym derivative process* $\Lambda_{\boldsymbol{\theta}} = \{\Lambda_{\boldsymbol{\theta}}(t),\ t \in [0,T]\}$, where

$$\Lambda_{\boldsymbol{\theta}}(t) = \exp\left\{ -\int_0^t \boldsymbol{\theta}_s^\top\, d\boldsymbol{W}_s - \frac{1}{2} \int_0^t \boldsymbol{\theta}_s^\top\, \boldsymbol{\theta}_s\, ds \right\} < \infty \tag{2.7.40}$$

almost surely for $t \in [0,T]$, is an $(\underline{\mathcal{A}}, P)$-*martingale*. By the Itô formula (1.5.8) it follows from (2.7.40) that

$$\Lambda_{\boldsymbol{\theta}}(t) = 1 - \sum_{i=1}^m \int_0^t \Lambda_{\boldsymbol{\theta}}(s)\, \theta_s^i\, dW_s^i \tag{2.7.41}$$

for $t \in [0,T]$. Since $\Lambda_{\boldsymbol{\theta}}$ is assumed here to be an $(\underline{\mathcal{A}}, P)$-martingale we have for $t \in [0,T]$

$$E\left(\Lambda_{\boldsymbol{\theta}}(t) \,\middle|\, \mathcal{A}_0\right) = \Lambda_{\boldsymbol{\theta}}(0) = 1. \tag{2.7.42}$$

We can now define a measure $P_{\boldsymbol{\theta}}$ via the Radon-Nikodym derivative

$$\frac{dP_{\boldsymbol{\theta}}}{dP} = \Lambda_{\boldsymbol{\theta}}(T) \tag{2.7.43}$$

by setting

$$P_{\boldsymbol{\theta}}(A) = E(\Lambda_{\boldsymbol{\theta}}(T)\, \mathbf{1}_A) = E_{\boldsymbol{\theta}}(\mathbf{1}_A) \tag{2.7.44}$$

for $A \in \mathcal{A}_T$. Here $\mathbf{1}_A$ is the indicator function for A and $E_{\boldsymbol{\theta}}$ means expectation with respect to $P_{\boldsymbol{\theta}}$.

Note that $P_{\boldsymbol{\theta}}$ is not just a measure but also a probability measure because

$$P_{\boldsymbol{\theta}}(\Omega) = E(\Lambda_{\boldsymbol{\theta}}(T)) = E\left(\Lambda_{\boldsymbol{\theta}}(T) \,\middle|\, \mathcal{A}_0\right) = \Lambda_{\boldsymbol{\theta}}(0) = 1 \tag{2.7.45}$$

as a result of the martingale property of $\Lambda_{\boldsymbol{\theta}}$.

Bayes Rule

It is useful to be able to change the probability measure when taking conditional expectations. The following *Bayes rule* establishes a relationship between conditional expectations with respect to different equivalent probability measures.

Assume for an equivalent probability measure $P_{\boldsymbol{\theta}}$ that the corresponding strictly positive Radon-Nikodym derivative process $\Lambda_{\boldsymbol{\theta}}$ is an $(\underline{\mathcal{A}}, P)$-martingale. Then for any given stopping time $\tau \in [0, T]$ and any $\mathcal{A}_\tau$-measurable random variable Y, satisfying the integrability condition

$$E_{\boldsymbol{\theta}}(|Y|) < \infty, \tag{2.7.46}$$

one can apply the *Bayes rule*

$$E_{\boldsymbol{\theta}}\left(Y \,\middle|\, \mathcal{A}_s\right) = \frac{E\left(\Lambda_{\boldsymbol{\theta}}(\tau)\, Y \,\middle|\, \mathcal{A}_s\right)}{E\left(\Lambda_{\boldsymbol{\theta}}(\tau) \,\middle|\, \mathcal{A}_s\right)} \tag{2.7.47}$$

for $s \in [0, \tau]$. This formula then allows us to change an expectation with respect to the real world probability to one with respect to, say, the risk neutral probability, under appropriate conditions as discussed in Sect. 3.3.

Girsanov Transformation

The following Girsanov transformation allows us to perform a measure transformation, which transforms a drifted Wiener process into a Wiener process under a new probability measure $P_{\boldsymbol{\theta}}$. More precisely, if for $T \in (0, \infty)$ a given strictly positive Radon-Nikodym derivative process $\Lambda_{\boldsymbol{\theta}}$ is an $(\underline{\mathcal{A}}, P)$-martingale, then the m-dimensional process $\boldsymbol{W}_{\boldsymbol{\theta}} = \{\boldsymbol{W}_{\boldsymbol{\theta}}(t),\, t \in [0, T]\}$, given by

$$\boldsymbol{W}_{\boldsymbol{\theta}}(t) = \boldsymbol{W}_t + \int_0^t \boldsymbol{\theta}_s \, ds \tag{2.7.48}$$

for all $t \in [0, T]$, is an m-dimensional standard Wiener process on the filtered probability space $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P_{\boldsymbol{\theta}})$.

Note that certain assumptions need to be satisfied before one can apply the above Girsanov transformation. The key assumption is that $\Lambda_{\boldsymbol{\theta}}$ must be a strictly positive $(\underline{\mathcal{A}}, P)$-martingale. For instance, if the Radon-Nikodym derivative process is only a strictly positive local martingale, then this does not guarantee that $P_{\boldsymbol{\theta}}$ is a probability measure, see Platen & Heath (2006).

A sufficient condition for the Radon-Nikodym derivative process $\Lambda_{\boldsymbol{\theta}}$ to be an $(\underline{\mathcal{A}}, P)$-martingale is the *Novikov condition*, see Novikov (1972a), which requires that

$$E\left(\exp\left\{\frac{1}{2}\int_0^T \boldsymbol{\theta}_s^\top \boldsymbol{\theta}_s \, ds\right\}\right) < \infty. \tag{2.7.49}$$

This condition is fulfilled, for instance, for the Black-Scholes model. However, it is not satisfied for the minimal market model, which will be outlined in Sect. 3.6.

Black-Scholes Formula

The following famous Black-Scholes pricing formula for a European call option with geometric Brownian motion as the dynamics for the underlying risky security can be obtained in various ways. These include applications of the Feynman-Kac formula, the Bayes rule and the Girsanov transformation.

Let the underlying risky security S_t satisfy the SDE

$$dS_t = r_t \, S_t \, dt + \sigma_t \, S_t \, dW_\theta(t) \tag{2.7.50}$$

for $t \in [0, T]$ with initial value $S_0 > 0$. Here W_θ denotes a standard Wiener process under the risk neutral probability measure P_θ. The deterministic functions of time r_t and σ_t denote the interest rate and volatility, respectively. Furthermore, in this simple two-asset market there exists a savings account B_t, which accrues the deterministic interest rate such that

$$dB_t = r_t \, B_t \, dt \tag{2.7.51}$$

for $t \in [0, T]$ and $B_0 = 1$. The European call payoff at maturity $T \in (0, \infty)$ is of the form

$$H = (S_T - K)^+ \tag{2.7.52}$$

with $K > 0$ denoting the strike price.

The pricing function $V(t, S_t)$ satisfies, by the risk neutral pricing formula, see (3.3.7), the conditional expectation

$$V(t, S_t) = E_\theta\left(\frac{(S_T - K)^+}{B_T}\,\middle|\,\mathcal{A}_t\right) \tag{2.7.53}$$

for $t \in [0, T]$, where E_θ denotes the expectation under the risk neutral probability measure P_θ with W_θ denoting a standard Wiener process under P_θ. The process W_θ would be a drifted Wiener process under the real world probability, see (2.7.48). Via the Feynman-Kac formula (2.7.6)–(2.7.8) it is straightforward to confirm that the pricing function $V(t, S)$ satisfies the following Black-Scholes PDE

$$\frac{\partial V(t,S)}{\partial t} + r_t \, S \, \frac{\partial V(t,S)}{\partial S} + \frac{1}{2}\,\sigma_t^2 \, S^2 \, \frac{\partial^2 V(t,S)}{\partial S^2} - r_t \, V(t,S) = 0 \tag{2.7.54}$$

for $t \in (0,T)$ and $S \in (0,\infty)$, with terminal condition

$$V(T,S) = (S-K)^+ \tag{2.7.55}$$

for $S \in (0,\infty)$. Furthermore, it can be shown that the option pricing function $V(t,S_t)$ satisfies the seminal *Black-Scholes formula*

$$V(t,S_t) = S_t\, N(d_1(t)) - K\,\frac{B_t}{B_T}\, N(d_2(t)) \tag{2.7.56}$$

with

$$d_1(t) = \frac{\ln\left(\frac{S_t}{K}\right) + \int_t^T \left(r_s + \frac{1}{2}\,\sigma_s^2\right)\, ds}{\sqrt{\int_t^T \sigma_s^2\, ds}}$$

and

$$d_2(t) = d_1(t) - \sqrt{\int_t^T \sigma_s^2\, ds}$$

for $t \in [0,T)$, see Black & Scholes (1973). Here $N(\cdot)$ denotes the standard Gaussian distribution function.

2.8 Exercises

2.1. Derive the joint transition density for two correlated drifted Wiener processes with constant correlation $\varrho \in (-1,1)$ and constant drifts $\mu_1, \mu_2 \in \Re$.

2.2. For the two-dimensional process in Exercise 2.1, write down a corresponding SDE driven by two independent Wiener processes W^1 and W^2.

2.3. Write down the transition density of a two-dimensional geometric Brownian motion that is a martingale, where the driving Wiener processes are correlated with parameter $\varrho \in (-1,1)$ and each of them having the volatility $b > 0$.

2.4. For the two-dimensional geometric Brownian motion in Exercise 2.3 write down the SDEs for its components driven by two independent Wiener processes W^1 and W^2.

2.5. Consider the following Merton model SDE with jumps

$$dX_t = a\, X_t\, dt + b\, X_t\, dW_t + c\, X_{t-}\, dN_t,$$

for $t \geq 0$ with $X_0 > 0$, where W is a standard Wiener process independent of the Poisson process N, which has intensity $\lambda > 0$. Verify the Feynman-Kac formula for the functional

$$u(t,x) = E\left(H(X_T) \,\middle|\, X_t = x\right)$$

for $(t,x) \in [0,T] \times \Re^+$.

2.6. Derive the first moment of a square root process with constant parameters $c > 0$, $b < 0$ and dimension $\delta > 2$ satisfying the SDE

$$dY_t = \left(\frac{\delta}{4}\, c^2 + b\, Y_t\right) dt + c\, \sqrt{Y_t}\, dW_t$$

for $t \in [0, \infty)$ and $Y_0 > 0$, where W is a Wiener process.

2.7. Prove that the ARCH diffusion model for squared volatility

$$d|\theta_t|^2 = \kappa\, (\bar{\theta}^2 - |\theta_t|^2)\, dt + \gamma\, |\theta_t|^2\, dW_t$$

has an inverse gamma density as stationary density.

2.8. Show that the squared volatility of the model

$$d|\theta_t|^2 = \kappa\, |\theta_t|^2\, (\bar{\theta}^2 - |\theta_t|^2)\, dt + \gamma\, |\theta_t|^3\, dW_t$$

has an inverse gamma density.

2.9. Compute the stationary density for the squared volatility for the Heston model

$$d|\theta_t|^2 = \kappa\, (\bar{\theta}^2 - |\theta_t|^2)\, dt + \gamma\, |\theta_t|\, dW_t.$$

http://www.springer.com/978-3-642-12057-2

Numerical Solution of Stochastic Differential Equations with Jumps in Finance

Platen, E.; Bruti-Liberati, N.

2010, XXVIII, 856 p., Hardcover
ISBN: 978-3-642-12057-2